

République Algérienne Démocratique et Populaire.  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique.

Université Saad Dahlab, Blida  
USDB.

Faculté des sciences.  
Département informatique.



**Mémoire pour l'obtention  
du diplôme d'ingénieur d'état en informatique.**

Option : Intelligence Artificielle

Sujet :

Détection de pathologies dans les images cytologiques par  
morphologie mathématique et classification connexionniste

Présenté par:

M<sup>elle</sup> Mennad Sakina et M<sup>elle</sup> Amia Naima

Dirigé par:

M<sup>elle</sup> REGUIEG F/Zohra et M<sup>elle</sup> BENBLIDIA Nadja

Novembre 2007

MIG-004-199-1

# Remerciements

*Nous remercions avant tout le bon dieu qui nous a aidé à réaliser ce*

*Modeste travail.*

*Nous tenons à remercier Melle. Reguieg Zora et Melle Benblidia Nadjia*

*Pour nous avoir encadrées et pour leur aide.*

*Nous remercions Monsieur le chef du département d'Informatique, tous  
les enseignants de la faculté des sciences de BLIDA et surtout ceux du*

*Département d'informatique.*

*Nous remercions les membres du jury pour nous avoir fait l'honneur de*

*Juger notre travail.*

*Nous remercions, de tout cœur, tous ceux qui ont contribué de près ou de*

*Loin à la réalisation de ce travail.*

*AMIA NAIMA L*

# Dédicace

*Que le salut de dieu soit sur vous, vous à qui je dédie affectueusement cet ouvrage.*

*Permettez moi d'énoncer les noms de quelques personnes qui me sont chers pour leur contribution dans cet ouvrage.*

*Je commencerai par mes très chers parents—mon cher Père, à qui je dédie tendrement ce mémoire, avec tout mon respect et ma tendre affection pour son grand cœur, ainsi qu'à ma mère unique au monde dans son genre pour son courage pour sa totale participation à la concrétisation de ce mémoire et son esprit de sacrifice durant tout mon apprentissage et son dévouement ; elle a été parfaite ; je lui dois ce que je suis et Dieu la récompense.*

*Je dédie ce mémoire à :*

- *mes sœurs Souad, Neelmine.*
- *mon frère Mohamed J. et ses amies.*
- *Ma grande famille chère par son nom et surtout mes grands parents paternels et mes grands maternels.*
- *mon binome Sakina et toute sa famille.*
- *toutes mes amies de l'université.*

*A. NAÏMA*

# Dédicace

*Je dédie ce mémoire à :*

*Mes très chers parents que je remercie infiniment pour tous leurs sacrifices, encouragements, conseils, et surtout leur affection et amour.*

*Mes sœurs Wissam, Houda.*

*Mes frères Nassim, Saïd.*

*Ma belle mère.*

*Ma grande famille chacun avec son nom*

*A. Naima et sa famille.*

*Une dédicace spéciale pour mon mari Kamel qui m'a aidé durant toutes mes études.*

*A toute la promotion d'informatique 2007 et toute mes amies de l'université.*

*M.SAKINA*

## Résumé

Nous cherchons à réaliser dans le cadre de ce travail, un système semi-automatique d'aide au diagnostic pour le cytopathologiste, suivant une méthode morphologique de segmentation d'images 2D couleurs de cytologie qui se base sur la ligne de partage des eaux améliorée. Celle-ci utilise une fonction de potentiel couleur combinant l'information locale et globale. Notre système permet de caractériser les cellules suivant des mesures quantitatives établies par rapport aux critères cytomorphologiques, dans le but d'une classification connexionniste suivant le modèle du perceptron multicouches, nous permettant ainsi, une aide à l'interprétation cytopathologique.

**Mot clés :** image cytologique, segmentation, morphologie mathématique, analyse quantitative, classification connexionniste.

## Abstract

We fetch to realize in this framework, a semi-automated computer-based system help of diagnosis. A morphological method for the color segmentation of cytological images is presented in 2D. This method is mainly based on watershed improved whose potential function blend local and global informations. Our system allows to characteristic of serous according to measurs quantified recording from criterias "cytomorphological", in the aim of a connected classification according the model of multi-layers perceptron, permitting to help the cytopathologist in his interpretation.

**Keywords :** cytological image, segmentation, Mathematic Morphology, quantified analysis, connected classification.

# Table des matières

Introduction Générale .....	01
<b>Chapitre I : Vision couleur et notions cytologiques</b>	
I. De la lumière aux couleurs .....	04
I.1 La lumière.....	04
I.2 La colorimétrie.....	05
II. Les modèles de la vision humaine et les espaces couleurs.....	05
II.1 Les espaces couleurs.....	05
II.1.1. L'espace RVB.....	06
II.1.2. L'espace CIE XYZ.....	06
II.1.3 L'espace CIE L*a*b* .....	06
II.1.4 l'espace L*U*V* .....	07
II.1.5. L'espace de couleur HSL (Hue, Saturation, Luminance).....	08
II.2 Ecart de couleur.....	08
III. Introduction à la cytologie.....	08
III.1 Anatomie pathologique et cytopathologie.....	09
III.2 Structure anatomique et cytologique.....	09
III.3 Types cellulaires en cytologie .....	10
III.4 Les prélèvements en anatomie et cytologie pathologiques.....	12
III. 4.1 Biopsie.....	12
III. 4.2 Examen extemporané.....	12
III. 4.3 Pièce opératoire.....	13
III.5 Les Colorations .....	13
III.6 Le screening .....	14
III.6.1 Le screening manuel.....	14
III.6.2 Le screening automatique.....	15
III.6.3 Structure d'un système informatique de rescreening.....	15
IV. Conclusion.....	15
<b>Chapitre II: Etat de l'art et segmentation par la LPE améliorée</b>	
I. Introduction.....	16
II. Segmentation par analyse des propriétés spatiales.....	18
II.1. Segmentation par division de régions.....	18
II.2 .Segmentation par croissance de régions.....	18
II.3. Segmentation par fusion de régions.....	19
II.4. Segmentation par division-fusion.....	19
II.5 Discussion.....	19
III. Segmentation par morphologie mathématique.....	19
III.1.Concepts premier pour la morphologie mathématique.....	20

III.1.1.	Erosion et dilatation.....	21
III.1.2.	Ouverture et fermeture.....	22
III.2.	Principales méthodes de segmentation par morphologie mathématique.....	23
III.2.1.	Segmentation par la L.P.E classique.....	24
III.2.2.	Segmentation par la L.P.E améliorée.....	24
III.3.	Conclusion.....	24
IV	Méthode de segmentation couleur adoptée.....	25
IV .1.	La ligne des partages des eaux.....	25
IV .1.1.	Fonction de potentiel couleur.....	25
IV.1.2.	Méthodologie d'utilisation.....	25
IV.1.4.	Choix de l'espace couleur.....	26
IV.1.5.	Choix du gradient couleur.....	26
IV.1.6.	Choix de la métrique.....	26
IV.1.7.	Choix de marqueurs.....	27
IV.1.8.	Choix de $\alpha$ .....	30
IV.2.	Classification de pixel par analyse d'histogramme2D.....	30
IV. 2.1.	L'espace couleur de l'histogramme.....	30
IV 2.2.	Espace morphologique non supervisé.....	30
IV 2.3.	Simplification d'image par lissage non linéaire d'un graphe.....	31
IV 2.4.	Hierarchie de partitions par fusion de régions.....	32
IV 2.5.	Amélioration de la L.P.E.....	33
IV. 3.	Conclusion.....	34

## Chapitre III : Caractérisation morphologique et classification connexionniste des images cytologiques

I	Caractérisation morphologique de la cellule.....	35
I. 1.	Description des paramètres morphologiques.....	35
I.2.	Rapport nucléo-cytoplasmique.....	36
I.2.1.	Création de l'image de noyaux.....	36
I.2.2	Création de l'image des cytoplasmes.....	38
I.3	Anisocaryose.....	39
I .3.1.	Repérage de noyaux.....	39
I.3.2.	Calculs intermédiaires sur chaque noyau repéré.....	40
I.4.	Difformité nucléaire.....	41
I.5.	Hyperchromatisme nucléaire.....	41
II.	Classification connexionniste.....	41
II.1.	Qu'est-ce qu'un réseau de neurones ?.....	41
II.2.	Modélisation d'un neurone.....	42
II.3.	Connectivité.....	43
II.4.	Apprentissage supervisé / non supervisé.....	43
II.5.	Calcul des poids synaptiques.....	43
II.6.	Quelques réseaux célèbres.....	43

II. 6.1. Le perceptron.....	43
II. 6.2. Les perceptrons multicouches (PMC).....	44
II. 6.3. Les réseaux de Hopfield.....	44
II. 6.4 .Les réseaux de Kohonen.....	44
II.7. Le modèle perceptron.....	45
II.7.1. Notion de perceptron.....	45
II.7.1.1. Le modèle.....	45
II.7.1.2. L'utilisation.....	45
II.7.1.3. Les limites de ce modèle.....	46
II.8. Classification par le réseau PMC.....	46.
II.8.1. Création de la base d'apprentissage.....	47
II.8.2. Le réseau de neurone implémenté.....	50
III. Conclusion.....	51

## **Chapitre IV : Implémentation et résultats**

I. Environnement de travail.....	52
I.1 Matériel utilisé.....	52
I.2 Description des images .....	52
I.3 Langage utilisé.....	52
I.4 Les avantages de JAVA.....	52
I.5 Liste des systèmes disposant d'une implémentation de Java.....	52
I.6 Environnement de Java.....	52
II. Présentation du logiciel cytomorph.....	53
II. 1 Interface du logiciel.....	53
II.2. Les menus.....	54
II.2.1. Menu « File ».....	54
II.2.2. Menu « Morphologie ».....	55
II.2.3. Menu « Segmentation ».....	55
II.2.4. Menu « Recognition ».....	57
II.2.5. Menu « Window ».....	57
II.2.6. Menu « Help ».....	57
II.3. La barre d'outils.....	58
III. Présentation des résultats et discussion.....	59
IV. Conclusion.....	66

Conclusion générale.....	67
--------------------------	----



## Liste des tableaux

<b>Tableau 1</b> : Contraste apporté pour chaque espace couleur.....	26
<b>Tableau 2</b> : Classification de 16 images cellulaires à l'aide des quatre paramètres de malignité.....	64
<b>Tableau 3</b> : Comparaison des résultats entre deux types de classification.....	66

## Liste des figures

Figure 1 : Chaîne de la vision par ordinateur.....	04
Figure 2: Espace RVB en 3D.....	06
Figure 3 : Le modèle $L^*a^*b^*$ .....	07
Figure 4 : Schéma du système mixte de repérage des points de couleur.....	07
Figure 5 : Les principaux éléments de la cellule.....	10
Figure 6 : Exemples de cellules.....	11
Figure 7 : Image cellulaire qui a subit une coloration.....	13
Figure 8 : Schéma général des différentes catégories de segmentation.....	17
Figure 9 : Segmentation par division de régions.....	18
Figure 10 : Exemple de segmentation en 4 régions.....	19
Figure 11 : la transformée de tout ou de rien H.M.T.....	21
Figure 12 : Erosion par un élément structurant.....	22
Figure 13 : Dilatation par un élément structurant.....	22
Figure 14 : Ouverture par un élément structurant.....	23
Figure 15 : Fermeture par un élément structurant.....	23
Figure 16 : Segmentation par l'histogramme 2D dans l'espace couleur $L^*a^*, L^*b^*$ .....	24
Figure 17 : Une image de microscopie et les régions extraites manuellement.....	26
Figure 18 : un exemple de segmentation du cytoplasme des cellules.....	27
Figure 19: Synoptique de segmentation couleur du cytoplasme.....	27
Figure 20 : Noyaux accolés.....	28
Figure 21 : Noyaux en amas.....	28
Figure 22 : Synoptique de la segmentation couleur des noyaux.....	29
Figure 23 : Fonction distance sur des noyaux accolés en vue 3D.....	29
Figure 24: Noyaux après séparation.....	29
Figure 25: (a) Histogramme 3D dans l'espace couleur. $L^*a^*b^*$ .....	31
Figure 26 : Image simplifiée.....	31
Figure 27: Algorithme de hiérarchie de partitions par simplification et fusion de régions...	33
Figure 28 Segmentation après la L.P.E améliorée.....	33
Figure 29 : Les images des différents types des cellules (a, b, c).....	35
Figure 30 : Exemples d'images à traiter.....	36
Figure 31 : Etapes permettant d'extraire l'image des noyaux de l'image originale.....	37
Figure 32: Critère d'appartenance à un noyau et à un filament.....	37
Figure 33: Image des cytoplasmes.....	38
Figure 34 : Algorithme de calcul centre de gravité.....	39
Figure 35: Distance caractéristique $x$ non significative dans l'image de deux noyaux accolés $N_i$ et $N_i'$ .....	40
Figure 36 : Existence de l'hyperchromatisme nucléaire dans les cellules ci-dessus.....	41
Figure 37 : Schéma d'un neurone biologique.....	42
Figure 38 : Structure d'un neurone artificiel.....	42
Figure 39 : Schéma représentant les PMC.....	44
Figure 40 : Deux cas de séparation.....	46
Figure 41 : Schéma expliquant le déroulement de réseau PMC.....	47
Figure 42 : Représentation de fonction sigmoïde.....	47
Figure 43: Algorithme d'apprentissage de réseau PMC.....	49
Figure 44 : principe de classification par le PMC.....	50
Figure 45 : Synoptique de l'application cytomorph.....	53
Figure 46 : Fenêtre principale de cytomorph.....	54

<b>Figure 48</b> : Menu « Morphologie ».....	55
<b>Figure 49</b> : Menu « Segmentation ».....	55
<b>Figure 50</b> : Menu « Color Space ».....	55
<b>Figure 51</b> : Menu « Filtre ».....	56
<b>Figure 52</b> : Menu « LPE ».....	56
<b>Figure 53</b> : Menu « Detecte ».....	56
<b>Figure 54</b> : Menu « Recognition ».....	57
<b>Figure 55</b> : Menu « Window ».....	57
<b>Figure 56</b> : Menu « Help ».....	57
<b>Figure 57</b> : Barre d'outils.....	58
<b>Figure 57</b> : choix de seuil.....	59
<b>Figure 58</b> : choix de membres d'itération pour les fonction morphologique.....	59
<b>Figure 58</b> : Résultats des différents filtrages .....	59
<b>Figure 59</b> : Détection des cellules d'intérêt.....	60
<b>Figure 60</b> : Représentation de l'image originale(12.a).....	60
<b>Figure 61</b> : Segmentation cytoplasmique et nucléaire avec précision du centre de gravité...63	63
<b>Figure 62</b> : Résultats de calcul des paramètres pour l'image de la figure (58.a).....	63
<b>Figure 63</b> : Apprentissage connexionniste par PMC.....	65
<b>Figure 64</b> : Classification des images par le réseau PMC.....	65



# *INTRODUCTION*

L'histoire de l'image est très ancienne : elle apparaît dès que l'homme commence à voir et à imaginer. C'est ainsi que, l'humain commence à dessiner dans les cavernes, sur les arbres ; écrire sur les peaux d'animaux (sérigraphie), etc... Cet être pensant, s'est distingué du reste des créatures terrestres, en exhibant comme un besoin de modifier et surtout, d'améliorer son mode de vie. Cette stratégie propre au genre humain, n'aurait jamais pu se tenir debout face aux siècles innombrables si elle n'était posée sur un piédestal imbattable. Ce dernier n'est autre que la communication et la transmission du savoir et de l'information : cette information que l'homme pense, stocke et développe.

Ainsi au fil du temps, il s'est trouvé évident que l'information, pour être communiquée, stockée et développée, se doit d'être limpide, et par conséquent universelle. Pour cela, l'homme a relevé le défi en élaborant un système de communication à la portée de tous, qui répond à ses besoins, passant du mode calligraphique (impression sur papier) au mode sonore à distance (téléphone, radio), jusqu'au mode visuel (impression sur écran). C'est ainsi que l'informatique, s'est trouvée mêlée à cette fabuleuse épopée de la vision humaine pour atteindre la vision par ordinateur. Les outils de communication à travers la vision par ordinateur, 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'examen 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 *traitement et analyse d'images*, le laboratoire de recherche pour le développement des systèmes informatisés 'LRDSI' et le laboratoire du traitement de signal et de l'image 'IATSI' ; portent un intérêt particulier à l'imagerie médicale, dont l'analyse d'images cytologiques. Nous cherchons dans ce cadre, à poursuivre la réalisation d'un système informatique semi-automatique permettant la détection de cellules anormales sur une

lame de prélèvement cytologique basée sur la morphologie mathématique. Notre travail s'intègre dans cette heuristique, en se basant sur le développement d'une technique améliorée d'extraction de l'information pertinente. L'analyse quantitative des cellules faisant appel à des attributs spécifiques, permet d'affiner la classification de nos images.

A partir d'un problème précis de segmentation et classification d'images et d'une implantation effective dans ce domaine, nous tentons de dégager les principes aboutissant à la réalisation d'un système robuste de segmentation d'images. Deux principes essentiels ont retenu notre attention : spécifier les connaissances sur l'image et sur les ressources procédurales du système.

Le but de notre système, est de détecter les cellules anormales ou suspectes qui pourraient échapper à la vigilance du pathologiste. Notre système traite des images couleurs cytologiques sur lesquelles il doit reconnaître les cellules suspectes ou anormales. Celles-ci sont présentées au pathologiste afin de vérifier son diagnostic, le rendant ainsi plus fiable. Nous cherchons à réaliser dans le cadre de ce travail, un système d'aide au diagnostic pour le cytopathologiste, suivant une méthode morphologique de segmentation d'images 2D couleur de cytologie utilisant l'information couleur comme information *a priori*. Cette méthodologie utilise des mesures quantitatives et des informations *a priori* sur les objets à extraire par des pathologistes (critères cytomorphologiques) dans le but d'une classification connexionniste nous permettant une aide à l'interprétation.

Le système que nous cherchons à développer, est basé autour de trois étapes essentielles :

✓ **La segmentation** : elle consiste à extraire des objets à partir d'images couleurs suivant la morphologie mathématique. Cette méthode de segmentation est basée sur la ligne de partage des eaux améliorée utilisant une fonction de potentiel couleur combinant l'information locale et globale.

✓ **La caractérisation** : elle fournit des caractéristiques des objets précédemment extraits, suivant une analyse quantitative du noyau et du cytoplasme. 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 suivant un modèle de réseaux de neurones, le modèle du perceptron multicouches.

Afin de mieux comprendre et suivre notre travail, dans le **premier chapitre**, nous définissons les caractéristiques d'une couleur, puis nous analysons les modèles de la vision humaine et les espaces couleurs que l'on retrouve dans la littérature. Nous présentons dans une seconde partie du chapitre, une vue générale du domaine cytologique illustrée par des images et des schémas.

**Dans le second chapitre**, nous abordons brièvement quelques techniques de segmentation d'images en régions. Nous nous consacrons plus en détails à la morphologie mathématique en accentuant notre étude sur la ligne de partage des eaux améliorée qui utilise une fonction de potentiel couleur combinant informations locale et globale déterminée à partir de l'analyse de l'histogramme 2D

Dans **le troisième chapitre**, nous étudions les paramètres de caractérisation spécifiques à l'imagerie cytologique ainsi que la classification connexionniste de nos cellules.

Enfin dans **le quatrième chapitre**, nous présentons l'application développée ainsi que les résultats illustrés et discutés.



# *CHAPITRE I*



La vision artificielle ou la vision assistée par ordinateur (V.A.O) symbolise toute la chaîne visuelle, partant de l'image brute et allant jusqu'à l'interprétation de son contenu. Il existe donc un aspect décisionnel très important lors de la phase de compréhension, d'où l'interaction avec l'intelligence artificielle. Le traitement d'image, s'intègre dans ce processus comme un outil, il n'a aucun pouvoir décisionnel car, il se définit comme un ensemble de tâches destinées à extraire de l'image, des informations qualitatives et quantitatives qui sont transmises au module de niveau supérieur [PER 90].

Le capteur dépend du domaine d'application. Le prétraitement réalise les tâches d'amélioration de l'image. L'extraction des caractéristiques, a pour rôle de faire une description de l'image compatible avec les objets contenus dans la base de reconnaissances. Enfin, le système décisionnel réalise la comparaison entre les caractéristiques qui lui sont fournies en amont et le contenu de la base de connaissance proprement dite (cf. Figure 1).

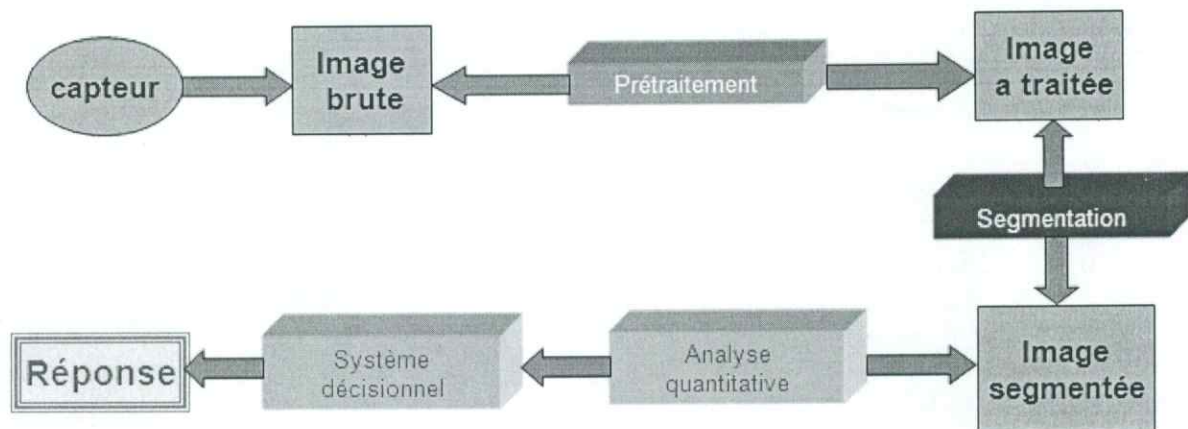


Figure 1 : Chaîne de la vision par ordinateur

Pour l'ordinateur, l'image est un ensemble de points (pixels) ayant chacun un niveau de gris, ou bien un niveau de couleur correspondant aux composantes; rouge, verte ou bleue, ou encore aux trois composantes, intensité, teinte et saturation.

## I. De la lumière aux couleurs

La couleur joue évidemment un rôle fondamental en informatique graphique et de façon générale, en théorie de la vision. Ce paragraphe rappelle donc quelques notions physiques sur la lumière et la couleur; pour exposer par la suite les moyens théoriques et numériques de coder une couleur. Plusieurs systèmes permettent de décrire les couleurs. Nous en présenterons cinq en précisant leurs avantages et inconvénients respectifs [PER 90].

### I.1 La lumière

La lumière visible est une énergie électromagnétique qui possède une distribution spectrale d'énergie. La lumière visible, est la zone du spectre comprise entre les longueurs d'onde 380 et 770 nanomètres (zone qui s'étend du violet au rouge, en passant par l'indigo, le bleu, le vert, le jaune et l'orange). Le spectre invisible est, quant à lui, constitué par les rayons gamma, X, ultraviolets, infrarouges et les ondes de radiofréquence [PER 90].

## I.2 La colorimétrie

La description d'une couleur, est une problématique complexe. On citera pour mémoire les travaux de la commission internationale de l'énergie (CIE), qui proposa en 1931[CIE31], un modèle peu intuitif basé, sur l'étude de la distribution spectrale de l'énergie lumineuse et la détermination de valeurs de trichromacités. Il est préférable donc, d'introduire le modèle de Munsell [PER 90], datant quant à lui de 1946, définissant une couleur par :

- ✓ **sa teinte**, qui est la notion utilisée pour distinguer les couleurs telles que le rouge, le vert, le jaune... Cette notion correspond en fait à celle de la longueur d'onde dominante de la distribution spectrale de la couleur.
- ✓ **sa saturation**, qui permet de mesurer la proportion de couleur pure par rapport au blanc. Cette notion permet de distinguer le rose du rouge, le bleu ciel du bleu outremer... Elle est liée à la largeur de bande du spectre de la couleur.
- ✓ **sa luminance**, qui est liée à la notion achromatique d'intensité d'excitation visuelle, est un facteur indépendant de la teinte et de la saturation.

## II. Les modèles de la vision humaine et les espaces couleurs

Un modèle de couleurs, décrit les couleurs que nous voyons et que nous utilisons dans les images numériques. Chaque modèle, RVB (Rouge, Vert, Bleu), CMJN (Cyan, Magenta, Jaune, Noir), (basé sur le, CMJN ou TSL [LEZ 02], représente une méthode différente de description des couleurs.

Un **espace colorimétrique**, est une variante d'un modèle de couleur caractérisée par sa propre **gamme** des couleurs. Par exemple, le modèle de couleurs RVB, recouvre plusieurs espaces colorimétriques : Adobe RVB, sRVB, ProPhoto RGB, etc. Chaque périphérique, a son propre espace colorimétrique et ne peut reproduire que les couleurs situées dans la gamme correspondante. Lorsqu'une image passe d'un périphérique à un autre, ses couleurs peuvent varier, car chaque périphérique interprète les valeurs RVB ou CMJN selon son propre espace colorimétrique [LEZ 02].

### II.1 Les espaces couleurs

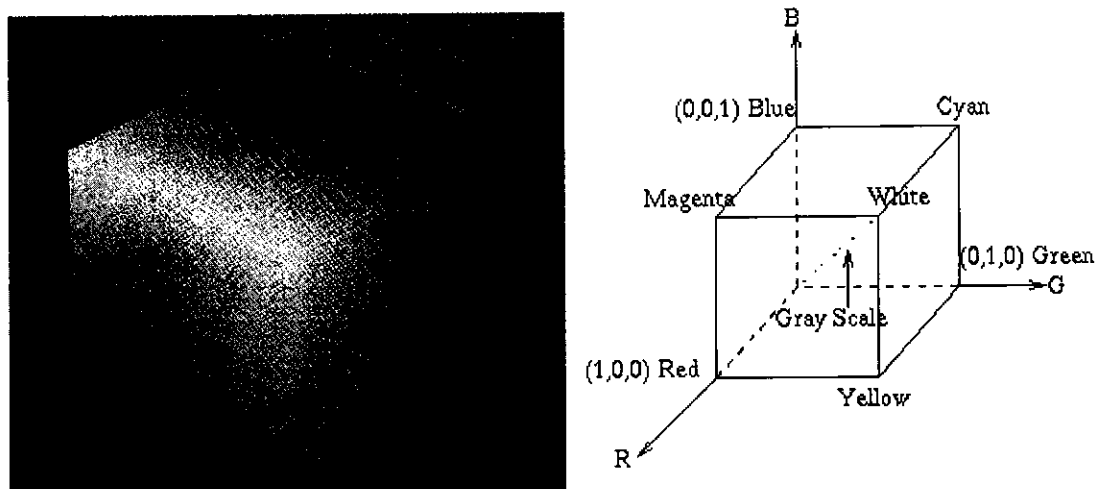
Une couleur est généralement représentée par trois composantes. Ces composantes définissent un espace des couleurs. On peut citer l'espace RVB, l'espace CIE XYZ ou YXY, ou encore l'espace Lab. Selon l'espace de couleurs choisi pour représenter une image couleur, le nuage des couleurs (c'est à dire l'ensemble des couleurs de l'image) n'aura pas la même répartition dans l'espace 3.

Les espaces de couleurs classiques, tels que le RVB, CIE XYZ, ...etc, sont issus d'une approche purement physique [LEZ 02].

Dans le cas d'autres espaces de couleurs, tels que l'espace Lab, l'approche physique est corrigée selon des données de la vision humaine [LEZ 02].

### II.1.1 L'espace RVB

L'espace RVB est sans doute l'espace de couleurs le plus utilisé. Les systèmes de télévision, s'y appuient fortement. La représentation des couleurs dans cet espace (cf. Figure 2) donne un cube appelé **cube de Maxwell** [LEZ 02].



**Figure 2:** Espace RVB en 3D

Ce modèle utilise un système de coordonnées cartésiennes et définit un cube unité comme indiqué sur la figure 2. Chaque axe correspond à une couleur primaire : rouge, verte et bleue. Une couleur est donc spécifiée en indiquant les contributions de chaque couleur primaire. Ces couleurs sont additives, c'est-à-dire que les contributions de chaque primaire sont ajoutées pour créer le résultat.

### II.1.2 L'espace CIE XYZ

Si nous utilisons trois spectres visibles [LEZ 02], comme base de l'espace visuel humain, il existera toujours des spectres ayant des coordonnées négatives sur cette base. Si nous utilisons par exemple, l'espace ou les trois primaires sont visibles. Le problème se pose dès que l'on désire travailler en synthèse additive (en n'ajoutant que des quantités positives de couleurs). La CIE a donc conçu un espace de couleur basé sur trois primaires X, Y et Z non visibles.

La transformation de l'espace à l'espace peut s'interpréter comme un changement de base donné par la matrice suivante:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.166 & 0.125 & 0.003 \\ 0.060 & 0.327 & 0.005 \\ 0.000 & 0.004 & 0.46 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

### II.1.3 L'espace CIE L\*a\*b\*

Le système L\*a\*b\* est issu du CIE XYZ (cf. Figure 3). Il essaye de prendre en compte la réponse logarithmique de l'oeil. Il possède le grand avantage d'être uniforme. Il est très utile

dans le cas de mélanges de pigments, par exemple, pour l'industrie graphique ou du textile [LEZ 02].

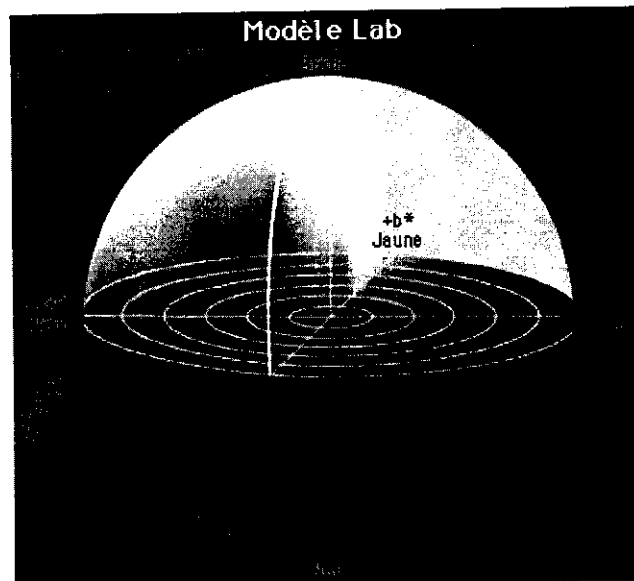


Figure 3 : Le modèle L\*a\*b\*

Une des difficultés majeures de ce système, est qu'il utilise un système mixte de repérage des points de couleur. La saturation est mesurée de manière cartésienne, alors que la teinte et la luminosité est mesurée de manière angulaire (cf. figure 4).

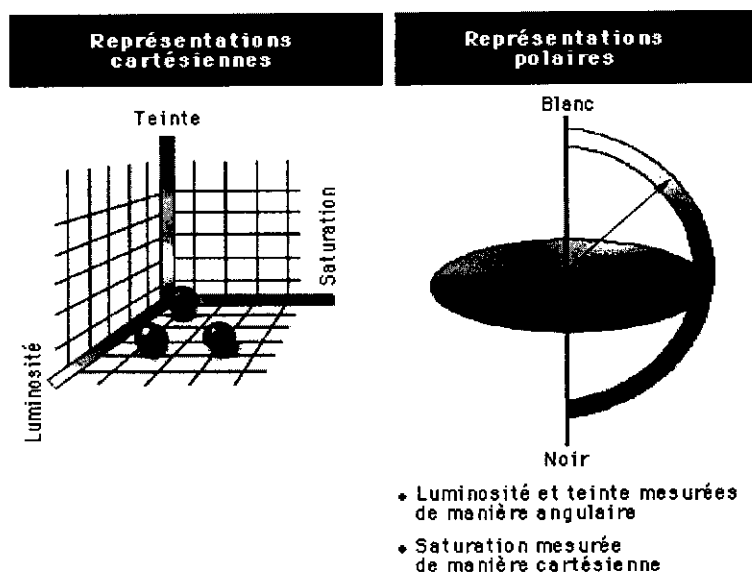


Figure 4 : Schéma du système mixte de repérage des points de couleur.

### II.1.4 l'espace 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 équidistances [LEZ 02]. 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 est perçue entre C et D. Les équations permettant de passer du système XYZ au  $l^*u^*v^*$  sont données ci-dessous :

$L^* = 116(Y/Y_0)^{1/3} - 16$	si $(Y/Y_0) > 0.008856$
$L^* = 9033(Y/Y_0)$	ailleurs
$U^* = 13L^*(U' - u')$	
$V^* = 13L^*(V' - v')$	
$U' = 4X/(X+15Y+3Z)$	$V' = 9X/(X+15Y+3Z)$
$u' = 4X_0/(X_0+15Y_0+3Z_0)$	$v' = 9Y_0/(X_0+15Y_0+3Z_0)$

### II.1.5. L'espace de couleur HSL (Hue, Saturation, Luminance)

Comme son nom l'indique il est basé sur les trois grandeurs de teinte, Saturation et Luminance. Cet espace est obtenu à partir de l'espace RVB. Dans cette espace, la teinte est singulière pour les points chromatiques ( $R=V=B$ ) et saturé, pour le noir ( $R=V=B=0$ ) [LEZ 02].

### II.2 Ecart de couleur

Lorsqu'il s'agit de travailler sur les couleurs d'une image, comme dans le cas d'une quantification, deux couleurs qui sont proches dans l'espace de couleur (au sens de la distance euclidienne le plus souvent), peuvent paraître assez différentes pour l'oeil, ce qui est le cas pour l'espace des couleurs RGB. Par contre, dans l'espace LAB qui est uniforme, deux couleurs proches en distance, le sont aussi pour l'oeil. C'est ici que réside l'intérêt de l'étude que nous présentons [LEZ 02].

## III. Introduction à la cytologie

La **biologie cellulaire**, ou **cytologie** est née avec l'invention du premier microscope optique (photonique) par Antonie van Leeuwenhoek [LEZ95] Il s'agit d'une discipline de la biologie, étudiant les cellules et leurs organites, les processus vitaux qui s'y déroulent ainsi que les mécanismes permettant leur survie (reproduction, métabolisme, homéostasie, néguentropie, communication) sans oublier la caractéristique principale de la cellule vivante, à savoir, la mort, qui peut être programmée génétiquement (apoptose) ou être le résultat d'une agression (nécrose).

L'histologie [LEZ95], est quant à elle, l'étude des cellules à un niveau supérieur, c'est-à-dire de leurs agencements en tissus et de leurs interactions (jonctions cellulaires simples, etc.).

La cytologie est une méthode d'étude de prélèvements cellulaires, faits à visée diagnostique pour laquelle la corrélation avec l'histologie et la clinique, a permis d'établir des règles de concordance diagnostique fiables.

### III.1 Anatomie pathologique et cytopathologie

L'anatomie pathologique (ou anatomo-pathologie) est la discipline médicale, qui permet la reconnaissance des anomalies des cellules et des tissus d'un organisme ; appelées lésions. Elle permet d'effectuer le diagnostic des maladies, de porter un pronostic et, plus généralement, en comprendre les causes et les mécanismes.

Elle s'appuie sur des techniques morphologiques (c'est à dire l'analyse de la forme des objets) : examen macroscopique (à l'oeil nu) et microscopique photonique (dit aussi « optique ») et électronique.

La cytopathologie étudie les cellules isolées obtenues, soit par frottis de **l'organe** à étudier (par exemple le frottis de dépistage gynécologique), soit par apposition d'un prélèvement, soit par aspiration d'un orifice naturel (par exemple aspiration bronchique), soit par ponction d'un liquide (pleural, péritonéal, liquide céphalo-rachidien...) ou d'une tumeur. Lorsque les cellules sont en suspension dans un liquide, elles sont centrifugées sur une lame (« cytocentrifugation ») [RAC 02].

#### ➤ Définition d'une cellule vivante

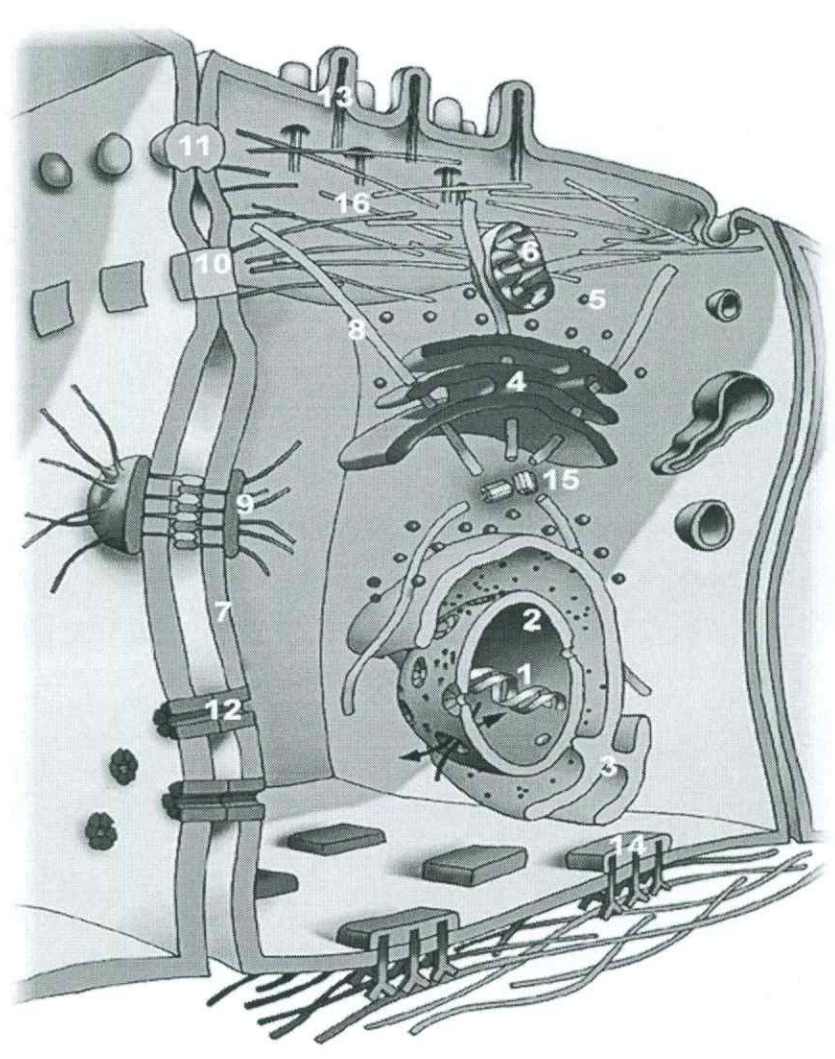
Une cellule se résume à six lois fondamentales : la cellule est la plus petite unité vivante. Tout être vivant est constitué de cellules; chacune d'entre elles, dérive d'une cellule préexistante, en général par division. Chaque cellule est limitée par une fine membrane plasmique qui la sépare du milieu environnant et des autres cellules. Toutes les cellules présentent de fortes similitudes biochimiques; elles sont de petite taille, de l'ordre du micromètre [RAC 02].

### III.2 Structure anatomique et cytologique

En 1838, Muller [Muel 38] 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. Des rapports sur les tumeurs en cytologie ont été réalisés ensuite par Beale [BEAL 60] sur des crachats, par Sanders [SAN 64] sur de l'urine par Dufour [DUF 04], sur des fluides cérébraux par Marrissi [MAR 09] sur des spécimens gastriques. En 1928, papanicolaou [PAP 47,54] présenta ses travaux sur l'étude cytologique du tractus 1940

On peut distinguer quatre niveaux structurant les cellules; les tissus, les organes et les appareils ou systèmes. A son niveau structural le plus simple, l'organisme est constitué de cellules: la cellule est l'unité fonctionnelle la plus petite de l'organisme. Les cellules peuvent être différentes entre elles selon leur fonction. Toutes les cellules vivantes possèdent des caractéristiques communes (cf. Figure 5).





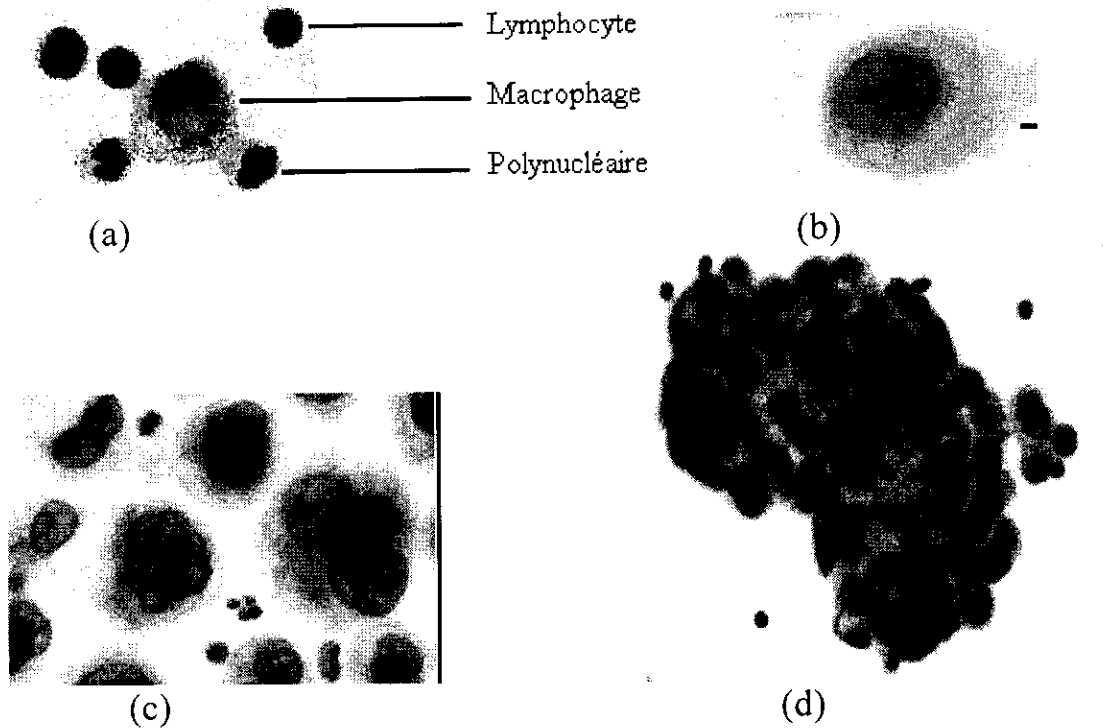
**Figure 5 :** Les principaux éléments de la cellule (Source: Institut Curie).  
 1. nucléole, 2. Enveloppe nucléaire, 3. Réticulum endoplasmique rugueux, 4. Appareil de Golgi, 5. Vésicules, 6. Mitochondrie, 7. Membrane plasmique, 8. Microtubules, 9. Neurone, 10. et 11. Jonctions adhérentes, 12. gap junction, 13. microvilosité, 14. desmosomes, 15. Centriole, 16. microfilaments

L'image ci-dessus présente une structure détaillée d'une cellule. Elle est composée de petits éléments appelés structures subcellulaires. Un composant fondamental de la cellule, est la membrane qui la délimite. Cette membrane entoure également de petits éléments appelés organites. La substance fluide où baignent ces organites, est appelée cytosol ou cytoplasme [RAC 02].

Le cytosol contient de nombreux enzymes essentiels et des métabolites. Le matériel génétique se situe dans le noyau. L'énergie nécessaire est fournie par la mitochondrie [RAC 02].

### III.3 Types cellulaires en cytologie

Nous remarquons en cytologie [LEZ95] plusieurs types de cellules (cf. Figure 6).



**Figure 6 :** Exemples de cellules (a) Lymphocyte active et non active, (b) cellule mésothéliale, (c) plusieurs cellules néoplasiques, (d) papille de cellule néoplasique.

Pour bien connaître les différents corps cellulaires, nous citons ici 6 types :

### 1-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 Curshamnn ou les calcosphérites,
- les cristaux de Charcot-leyden extra- cellulaire,
- les éléments mycéliens,
- les substances étrangères (grains de colorants, poussières, cellules végétales, grains de talc, dépôts de carbone des mines de crayon bullesd'air altermaria).
- 

### 2-Les cellules épithéliales normales

- *mésotéliales en bague à chaton,*
  - ✓ Mitoses
  - ✓ Binucléés
  - ✓ Multinucléés
- *Les malphigiennes.*
  - ✓ Intermediares
  - ✓ Superficielles



**3- Les cellules inflammatoires**

- Les polynucléaires,
- Neutrophiles,
- Eosiniphiles,
- Altérés (pyocytes),
- Les lymphocytes,
- Activés,
- Non activés,
- Les histiocytes,
- Macrophages.

**4- Les cellules tumorales primaires**

- Cellules du mésothéliome malin

**5- Les cellules tumorales secondaires**

- Cellules des lymphomes malins en cellules isolés,
- Sarcomes en cellules isolés,
- Les mélanomes malins.

On trouve aussi d'autres types de cellules, telles que :

- Les cellules épithéliales normales,
- Les cellules inflammatoires,
- Les cellules tumorales primaires.

**III.4 Les prélèvements en anatomie et cytologie pathologiques**

Un prélèvement est un acte médical [LEZ95] et il, ne doit pas être traumatisant pour les cellules. Les techniques anatomopathologiques sont appliquées à des prélèvements de tissus ou de cellules (cytopathologie) effectués, pendant la vie du malade (biopsie d'un tissu, examen d'une pièce opératoire...), ou après sa mort (autopsie, aussi appelée nécropsie). On peut savoir plusieurs types de prélèvements :

**III. 4.1 Biopsie**

La biopsie est le prélèvement d'un fragment de tissu effectué sur un être vivant. Il peut être réalisé au bistouri, à l'aiguille, à la griffe ou à l'emporte pièce (« punch »). La biopsie peut être effectuée à l'oeil nu, lors d'une intervention à ciel ouvert, ou sous endoscopie, par exemple. Lorsque l'ensemble de la lésion est prélevé, on parle de biopsie exérèse [LEZ95].

**III. 4.2 Examen extemporané**

Un examen extemporané [LEZ95], donne le résultat au clinicien en quelques minutes. Il est demandé lorsque son résultat est susceptible de modifier l'acte chirurgical (élargissement d'une exérèse, vérification de la qualité d'un greffon...).

### III. 4.3 Pièce opératoire

La résection [LEZ95] d'une pièce chirurgicale, doit être suivie d'un examen anatomopathologique destiné au bilan des lésions.

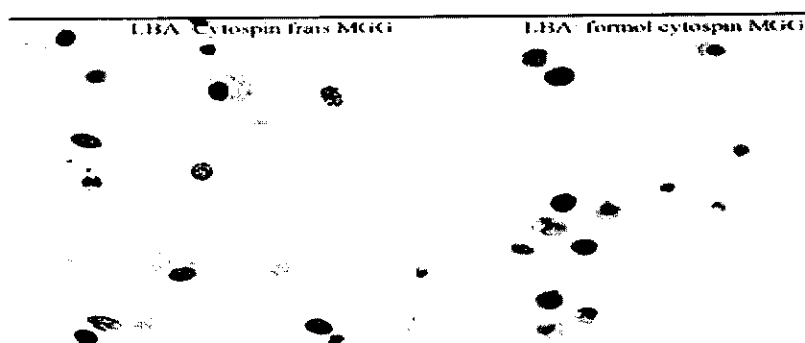
Parmi les prélèvements cytologiques, nous constatons :

- **Lames de cytologie** (cytoponction du sein ou d'organes pleins),
- **Liquides des séreuses** (plèvre, péritoine), du lavage bronchoalvéolaire (LBA), du liquide céphalorachidien (LCR), et du liquide articulaire.
- **Couleur:** (citrin, hématique, eau de roche).

### III.5 Les Colorations

Il n'existe pas de technique de coloration polyvalente [PAP 54]. Selon le but à atteindre, en l'occurrence selon la nature chimique des différents constituants cellulaires (cf. Figure 7) à mettre en évidence, on fera appel à des techniques différentes.

La coloration ne peut être réussie que si, elle est précédée d'une fixation appropriée visant à préserver l'intégrité des structures à colorer. Le principe de Fixation, est de Maintenir les structures pour éviter leur dégradation (digestion enzymatique, dégradation des cellules dont les composants sont dénaturés mais préservés de façon reproductible).



**Figure 7 :** Image cellulaire qui a subi une coloration

Les principales colorations cytologiques sont :

- Giemsa: hématologie
- Papanicolaou: fixation, frottis, liquides
- Agents pathogènes : Gram Weigert, Grocott , Ziehl
- Histochimiques: PAS, ORO, Perls
- Immunochimiques:
- Hybridation in situ:

### III.6 Le screening

L'analyse des prélèvements cytologiques se fait visuellement grâce à un cytopathologiste [LEZ 98]. Nous distinguons ce screening dans ce qui suit :

#### III.6.1 Le screening manuel

Une fois le prélèvement effectué, il est étalé sur une lame. Les cellules sont colorées par la coloration de papanicolaou et c'est ainsi que, la phase d'expertise peut commencer.

Celle-ci, consiste en une étape de lecture de la lame par un expert en cytologie (un cytotechnicien). Il repère les cellules intéressantes. 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 fiable.

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). Il inspecte une partie de la lame à travers un champ microscopique. 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 pour une détection plus fiable.

Le screening manuel des prélèvements cytologiques est donc difficile, fatiguant consommateur de temps et est décrit comme un travail intense, particulièrement complexe (le résultat repose sur le point de vue de l'être humain). Le faible nombre de cellules anormales (comparativement aux cellules normales) observées durant le screening, implique une grande concentration du cytotechnicien.

Un cytotechnicien observe entre 30 et 200 lames par jour. Sa culture médicale, est remise en cause chaque jour à cause de ce facteur subjectif inévitable. Effectivement, quelques erreurs peuvent apparaître et causer des faux négatifs. Les préoccupations principales en cytologie, sont celles qui permettent d'augmenter le contenu de vérité contenu dans le diagnostic. Un cytopathologiste s'attache donc, à réaliser les moins de faux positifs possibles mais surtout de faux négatifs.

Certains travaux alarmants rencontrés dans la littérature, font état de 50% de faux négatifs pour la cytologie du col [KOS 89]. Les principales causes d'erreurs en cytologie sont :

- les prélèvements non significatifs,
- 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 [REV 67], par arrêté fédéral en réponse à un contrôle de qualité dans les services de cytopathologie. Cependant après plusieurs années, il n'est pas apparu que ce genre de contrôle des lames élimine des erreurs de screening. La relecture rapide de toutes les lames donne de meilleurs résultats mais reste consommatrice de temps.

### III.6.2 Le screening automatique

Afin d'améliorer le contrôle de la qualité du screening, certaines recherches [LEZ 98], se sont orientées vers une automatisation totale du screening. La mise en place d'un système informatique, peut apporter au cytopathologie un plus dans la précision de l'analyse qualitative. En effet, l'examen des lames par un cytotechnicien est une tâche subjective (celle-ci n'est pas infaillible) ; un système informatique peut permettre de donner une analyse objective de lames de tels systèmes basés sur l'analyse d'images.

### III.6.3 structure d'un système informatique de rescreening

Les systèmes informatiques de rescreening [Ram 05], sont conçus pour opérer de façon postérieure au screening manuel. Cependant, ils peuvent éventuellement être utilisés pour fonctionner de manière antérieure. Effectivement, 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 les composantes cellulaires à partir d'images couleurs,
- **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, etc.), de façon à déterminer le type de la cellule à classer et plus précisément à détecter son caractère bénin ou malin.

## IV. Conclusion

Le traitement d'image est devenu une part essentielle de l'activité scientifique et technologique d'aujourd'hui. Ce chapitre, a été consacré aux notions générales de la vision humaine ; spécifiquement la colorimétrie et les espaces couleurs. Une étude globale sur le domaine cytopathologique, a été appréhendée afin de mieux analyser l'information. Le chapitre suivant, présente un aperçu des méthodes de segmentation en régions et plus en détails la morphologie mathématique.



## *CHAPITRE II*

## I. Introduction

La segmentation est une tâche très importante et difficile [1], car elle dépend de la nature de l'image et de l'objectif de l'analyse. La littérature [2], propose de nombreuses techniques de segmentation, mais la plupart d'entre elles ont besoin de plusieurs paramètres dont le réglage exige souvent l'expertise humaine. Dans ce qui suit, nous présentons quelques méthodes de segmentation, en choisissant deux chemins de conception qui se basent soit sur une analyse des propriétés spatiales où l'on utilise les différents algorithmes présentés pour extraire une région ; ou sur la morphologie mathématique, qui utilise l'information couleur comme information a priori.

Avant de rappeler les différentes méthodes de segmentation couleur (cf. Figure 8), il nous semble important de préciser en quoi celles-ci peuvent varier. Les méthodes de segmentation couleur présentent plusieurs particularités que l'on retrouve systématiquement [LEZ02] :

- **La méthode utilisée** : elle caractérise la manière d'extraire les régions d'une image grâce à son information couleur [LEZ 02].
  
- **Le choix de l'espace de travail** : les images couleurs sont généralement acquises et/ou stockées dans l'espace RVB [COC 95]. Les traitements peuvent être effectués dans cet espace. Cette manière de procéder, présente l'avantage de garantir l'intégrité des données, mais elle entraîne un certain nombre d'inconvénients :
  - ✓ Les composantes  $y$  sont souvent fortement corrélées,
  - ✓ il est difficile d'y interpréter la notion de couleur telle qu'elle est habituellement perçue par l'œil humain,
  - ✓ la différence perceptuelle entre deux couleurs ne coïncide pas avec leurs distances [LEZ 02].
  
- **La métrique utilisée** : La distance euclidienne et la distance de Mahalanobis [LEZ 98], sont les plus utilisées. Le problème de la mesure des différences colorimétriques, est un des plus importants problèmes qui se posent pour l'analyse de la distribution des couleurs.

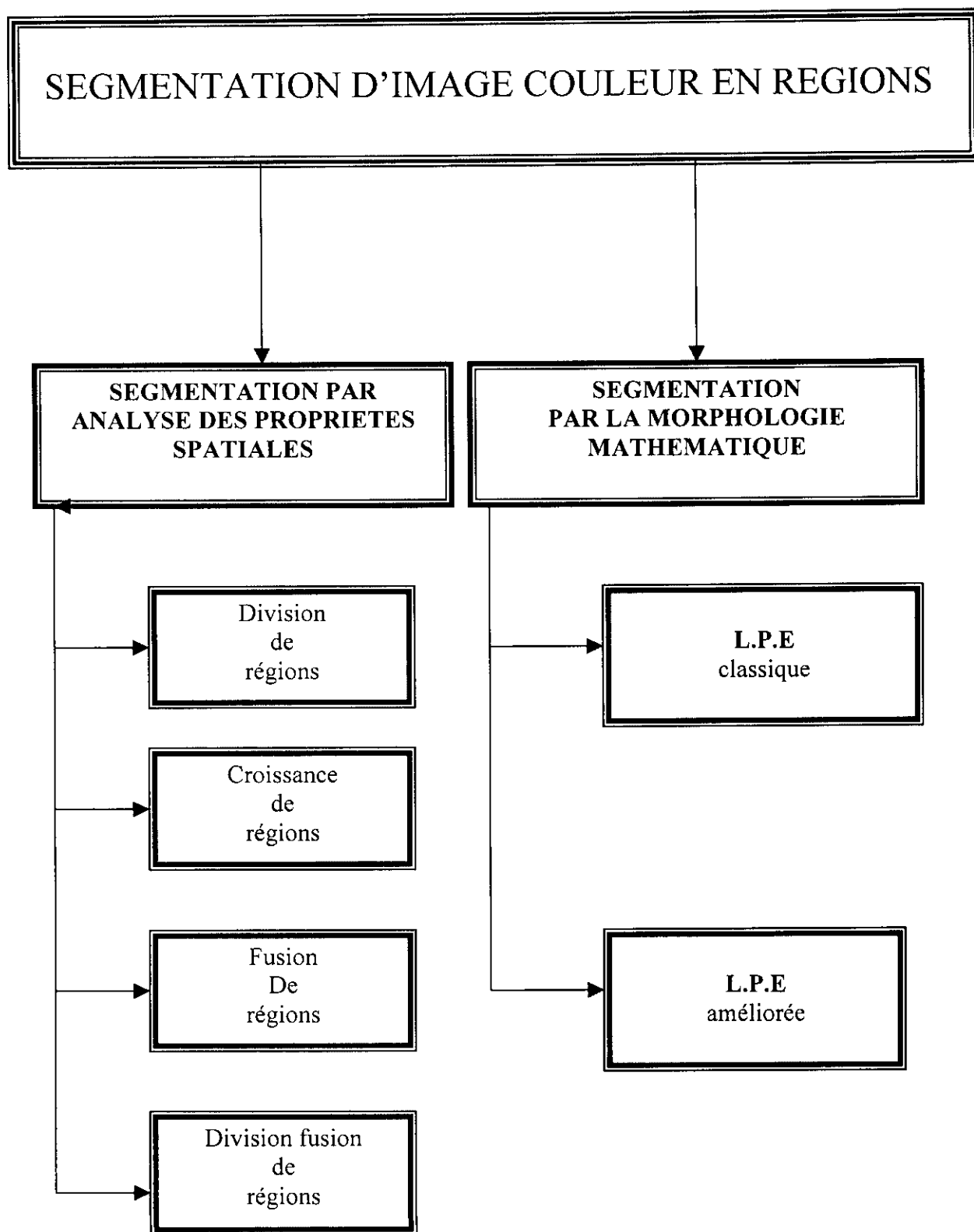


Figure 8 : Schéma général des différentes catégories de segmentation

## II. Segmentation par analyse des propriétés spatiales

La segmentation des images, consiste à regrouper les pixels de ces images qui partagent une même propriété pour former des régions connexes. Les régions sont déterminées en fonction de leurs propriétés intrinsèques (agrégation). Contrairement à l'extraction des contours, qui s'intéresse aux bords des régions, la segmentation en régions homogènes vise à segmenter l'image en se basant sur les propriétés intrinsèques des régions [LEG 02]. Nous distinguons ici quatre approches :

### II. 1 Segmentation par division de régions

Le principe de la segmentation par division [LEG 02] (split) (cf. Figure 9) est réalisé suivant :

- ✓ la définition d'un critère d'homogénéité : forme a priori de l'histogramme, extrema de l'image (valeurs minimum et maximum), valeurs identiques, variance limitée, approximation par un polynôme de degré N, etc.
- ✓ Le test de la validité du critère sur l'image. Ainsi, si le critère est valide, l'image est segmentée (arrêt de la méthode), sinon, l'image est découpée en zones plus petites et la méthode est réappliquée sur chacune des zones.

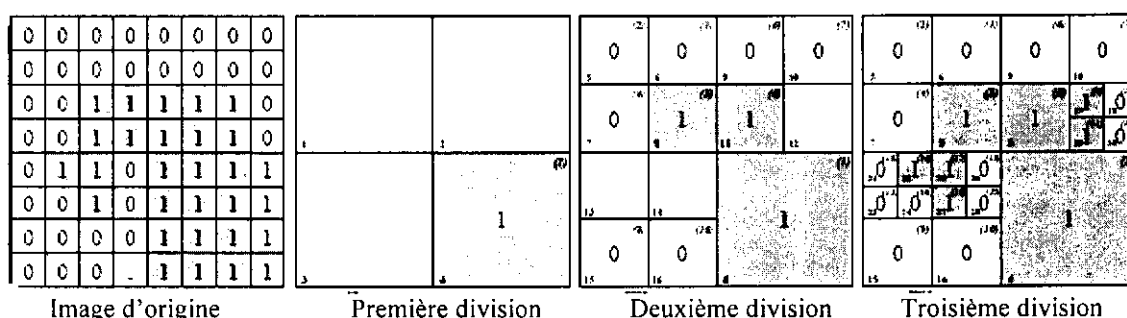


Figure 9 : Segmentation par division de régions

### II. 2 Segmentation par croissance de régions

La croissance de régions [3], est une méthode particulière de la segmentation. Elle vise à grouper les points adjacents dont les attributs varient d'une façon négligeable. D'un point de vue topologique, la segmentation par croissance de régions correspond à une partition de l'image I en L régions connexes  $R_i$ ,  $i=1, \dots, L$  telle que l'union des L supports de ces régions redonne le support de l'image originale. Le problème fondamental lié à la croissance de régions, est la définition de l'attribut A et du prédicat d'uniformité  $P(A, R_i)$  qui est vrai dans chaque région  $i=1, \dots, L$  et qui est faux pour l'ensemble de deux régions adjacentes  $R_i$  et  $R_j$ .

$$P [A, R_i] = \text{VRAI}, i=1, \dots, L.$$

$$P [A, R_i \cup R_j] = \text{FAUX}, i \neq j \text{ et } R_i \text{ adjacent à } R_j.$$

Un des points attractifs de la segmentation par croissance de régions, est lié au concept même de la région. En effet, plus une région est grande, plus la croissance statistique des propriétés de ses points est précise [KUNT 98].



### II. 3 Segmentation par fusion de régions

Une région est un ensemble de points n-connexes de même niveau de gris, qui seront 4-connexes dans l'exemple suivant (cf. Figure 10). La seule fonction requise, est de pouvoir énumérer les pixels qui composent une région. Il est utile pourtant de mémoriser, pour chaque région, le nombre de pixels, la somme de leurs valeurs (pour calculer la moyenne) et d'autres informations qui serviront à déterminer rapidement des critères de fusion. Lors de la fusion de deux régions, l'une est choisie arbitrairement pour "absorber" l'autre. Les données de la région absorbante sont alors "augmentées" de celles de la région absorbée.

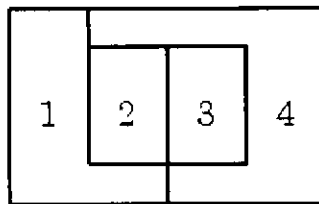


Figure 10 : Exemple de segmentation en 4 régions

Deux régions sont adjacentes, si au moins un pixel de l'une est voisin direct (pour la "4-connexité") d'un pixel de l'autre. Cela définit très naturellement une structure de graphe non orienté. La figure 8 donne un exemple, les adjacences sont alors  $(1 \leftrightarrow 2)$ ,  $(1 \leftrightarrow 4)$ ,  $(2 \leftrightarrow 3)$ ,  $(2 \leftrightarrow 4)$  et  $(3 \leftrightarrow 4)$ . La mise à jour des adjacences, lors d'une fusion constitue l'une des difficultés principales de ce sujet. On étudiera sur la figure 8, les conséquences d'une absorption de la région 4 par la région 1. Les adjacences restantes sont alors  $(1 \leftrightarrow 2)$ ,  $(1 \leftrightarrow 3)$  et  $(2 \leftrightarrow 3)$ . Le coût de mise à jour doit être linéaire dans le nombre d'adjacences concernées, c'est-à-dire, celles des deux régions à fusionner ainsi que celles de toutes leurs voisines directes [PHI 04].

### II. 4 Segmentation par division fusion

La segmentation par division fusion, regroupe les deux types d'approches précédentes. Tout d'abord, l'image est divisée en régions homogènes qui respectent des critères globaux, puis les régions adjacentes qui répondent à des critères locaux, sont fusionnées [1].

### II.5 Discussion

Les méthodes de segmentation en régions d'image couleurs précédentes, considèrent une région comme un ensemble de pixels connexes ayant une couleur proche d'une valeur moyenne. Ces méthodes analysent directement l'image matricielle et font intervenir l'information couleur multidimensionnelle dans la phase de décision.

## III. Segmentation par morphologie mathématique

La morphologie mathématique est une science de la forme et de la structure [BEU 90]. On a quatre définitions selon le domaine appliqué :

- **En mathématique :** Algèbre des treillis, géométrie ensembliste, intégrale, stériorologie et modèles topologico-probabilistes,

- **En physique** : Démarche fondée sur la théorie des ensemble destinée à lier les propriétés physiques des objets et texture des structures,
- **En traitement de signal** : Technique non linéaire de traitement du signal reposant sur des opérations de supérieur et d'inférieur.
- **En informatique** : c'est un algorithme et un programme permettant d'effectuer du traitement d'image sur ordinateur ou sur matériel spécialisé [SER 00].

Le principe de base de la morphologie mathématique, est de comparer les objets que l'on veut analyser à un autre objet de référence, de taille et de forme connues, appelé **élément structurant** [COS 89]. En quelque sorte, chaque élément structurant fait apparaître l'objet sous une nouvelle forme. Fondamentalement, la morphologie mathématique a un caractère ensembliste et ceci correspond à la démarche dont nous disposons pour analyser quantitativement une image [COS 89].

La structure d'une stratégie de segmentation par morphologie mathématique, se décrit par trois étapes séquentielles : la simplification, l'extraction de marqueurs, la croissance [LEZ 00].

- ✓ **La simplification** : La simplification de l'image, consiste en une phase de prétraitements dont le but est de simplifier les images. La simplification peut consister en un lissage de l'image, grâce à une régularisation réalisée par un filtre. Des opérateurs connexes basés, sur la reconstruction morphologique permettent d'éliminer certains détails ou objets selon des critères de taille, de forme, ou de couleur.
- ✓ **L'extraction** : L'extraction de marqueurs est l'étape d'initialisation de la croissance : c'est une étape orientée vers une extraction grossière des objets. Une connaissance a priori sur l'information des images à traiter et sur les caractéristiques des objets dans l'image, permet de choisir des opérateurs de traitement d'images. Ces opérateurs permettent, en exploitant les caractéristiques des objets, d'extraire partiellement ou totalement les objets sans toutefois être précis au niveau de la localisation. Pour les images couleurs de microscopie, la principale information a priori se situe au niveau de la couleur des objets. Pour cette raison, l'extraction des marqueurs se doit donc d'utiliser la couleur ou une composante couleur.
- ✓ **La croissance** : A partir des marqueurs obtenus à l'étape précédente, on utilise l'une des principales transformations de morphologie mathématique : la ligne de partage des eaux. Cette dernière réalise une croissance et permet d'obtenir les régions finales correspondant aux objets segmentés. Cette étape doit également faire appel aux informations a priori disponibles sur les objets et en particulier leur couleur. L'utilisation de la couleur dans le calcul de la ligne de partage des eaux, permettra d'obtenir des résultats qui correspondront mieux aux changements significatifs présents dans les images [LEZ 00].

### III.1 Concepts premiers pour la morphologie mathématique

La morphologie mathématique [SER 00] est basée sur des opérateurs principaux dont :

### III. 2.1 Erosion et dilatation

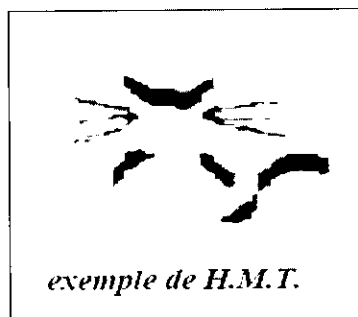
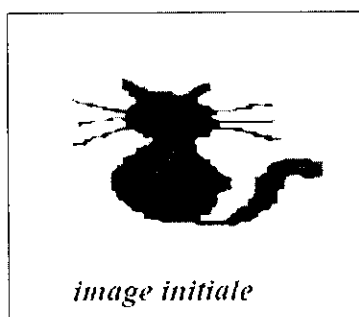
#### ➤ Transformée de tout ou de rien (H.M.T) et érosion ensembliste

Les objets d'étude sont ici les ensembles  $X \subseteq E$ . La morphologie mathématique les décrit, en associant à tout  $X \in E$  deux sondes  $A(x)$  et  $B(x)$  et en testant si l'une est dans  $X^c$  et l'autre dans  $X$ . Les fonctions  $x \rightarrow A(x)$  et  $x \rightarrow B(x)$  sont appelées **éléments structurants**, et le transformé par tout ou rien de  $X$  (H.M.T) se définit par l'opération (cf. Figure 11) :

$$\eta(X) = \{z: A(z) \subseteq X^c; B(z) \subseteq X\}$$

Quand  $A = \emptyset$ ,  $\eta(X)$  devient l'érodé de  $X$  par l'élément structurant (variable)  $B$  et s'écrit :

$$\varepsilon_B(X) = \{z: B(z) \subseteq X\}$$



**Figure 11** : la transformée de tout ou de rien H.M.T

L'**érosion ensembliste** (cf. Figure 12) est l'opération commutée avec  $\cap$  :

$$\begin{aligned} \varepsilon_B(\cap X_i) &= \{z: B(z) \subseteq \cap X_i\} \\ &= \cap \{z: B(z) \subseteq X_i\} \\ &= \cap \varepsilon_B(X_i) \end{aligned}$$

Il s'agit donc bien d'une **érosion**, et on appelle soustraction de Minkowski la relation suivante :

$$\varepsilon_B(X) = X \ominus B = \{z: B_z \subseteq X\}$$

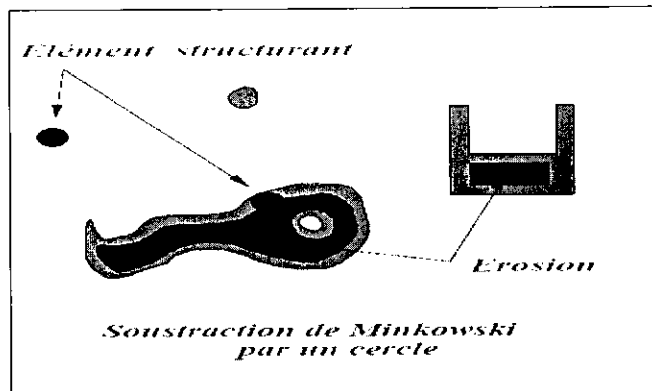


Figure 12 : Erosion par un élément structurant

➤ **Dilatation ensembliste**

Cette fois, les limites de l'opérateur sont contraintes à l'intérieur de A (sous-ensemble). Les équivalences :

$$X \subseteq \varepsilon_B(Y) \Leftrightarrow \{x \in X \Rightarrow B(x) \subseteq Y\} \Leftrightarrow \cup \{B(x), x \in X\} \subseteq Y$$

Conduisent à l'opération,

$$\delta_B(X) = \cup \{B(x), x \in X\}$$

qui commute avec l'union. C'est donc une dilatation (cf. Figure 13) dite adjointe de  $\varepsilon$ , et on appelle addition de Minkowski la relation suivante ;

$$\delta_B(X) = X \oplus B = \cup \{B(x), x \in X\}$$

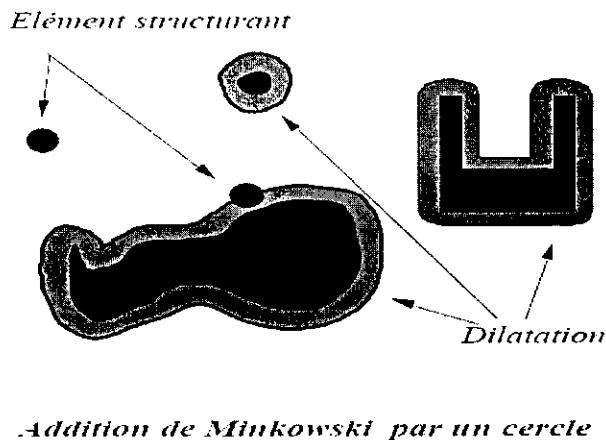


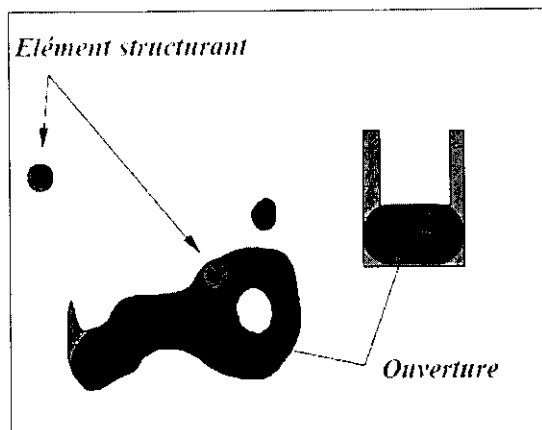
Figure 13 : Dilatation par un élément structurant

**III.2.2. Ouverture et fermeture**

Plusieurs ensembles différents, peuvent avoir le même érodé, ou le même dilaté. Néanmoins, entre tout les inverses possibles, il en existe un plus petit. On l'obtient en composant l'érosion par la dilatation adjointe [SER 00].

La relation ci-dessous, se nomme **ouverture** (cf. Figure 14) par adjonction que l'on note,

$$\gamma_B = \delta_B \varepsilon_B = X \circ B = [(X \ominus B) \oplus B]$$

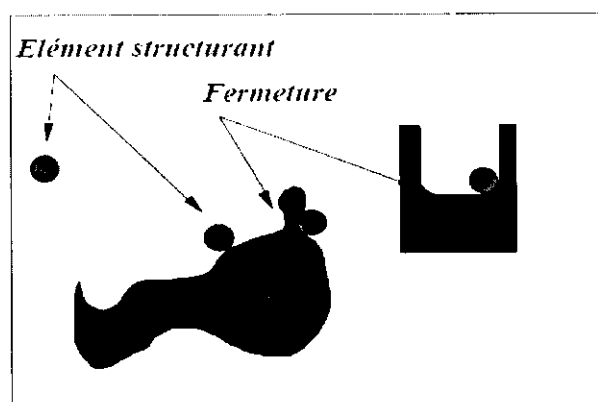


*L'ouverture élimine les petites îles , ouvre les isthmes étroits et réduit les caps .*

**Figure 14 :** Ouverture par un élément structurant

La relation ci-dessous, se nomme **fermeture** (cf. Figure 15) par adjonction que l'on note,

$$\varphi_B = \varepsilon_B \delta_B = X \bullet B = [X \oplus B] \ominus B]$$



*La fermeture bouche les petits lacs , ferme les chenaux et comble partiellement les golfes*

**Figure 15 :** Fermeture par un élément structurant

### III.2 Principales méthodes de segmentation par morphologie mathématique

En morphologie mathématique, l'une des principales transformations, est la ligne de partage des eaux. Celle-ci a été étendue au cas des images couleurs. L'innovation principale, est venue des travaux de Meyer [Mey92] qui a proposé un algorithme dont les similarités avec les méthodes de croissance de régions, sont nombreuses. Il utilise des différences de couleur entre pixels, ainsi qu'une structure de données de type file hiérarchique pour réaliser une croissance de régions rapide où les régions se propagent à partir des marqueurs. La mesure utilisée pour déterminer la probabilité d'appartenance d'un pixel à une région, et donc sa

probabilité d'agrégation, s'exprime par une fonction de potentiel. Cette fonction consiste en une mesure de similarité entre un pixel et une région voisine à laquelle il pourrait être agrégé. Les recherches qui ont été menées, se situent surtout au niveau de cette fonction de potentiel [LEZ 00].

### III. 2.1 Segmentation par la L.P.E classique

Si l'on considère l'image comme une surface dans l'espace de dimensions 3, on peut utiliser la terminologie de la géographie qui définit la ligne de partage des eaux comme la crête qui forme la limite entre deux bassins versants. Celle-ci s'obtient par amincissement morphologique effectué avec l'élément structurant  $L^2$  itéré jusqu'à stabilisation [COC 95].

### III. 2. 2 Segmentation par la L.P.E améliorée

La notion de base de segmentation par la L.P.E améliorée, est la même que la L.P.E classique mais cette fois ci, on utilise les méthodes de classification par analyse d'histogrammes (cf. Figure 16). Chaque pixel d'une image couleur  $I$ , est un vecteur à trois composantes. L'histogramme complet de l'image est donc, un histogramme tridimensionnel (3-D). On peut en déduire trois histogrammes 2-D, par projection de l'histogramme 3-D sur 2 des 3 plans colorimétriques de l'image. Un histogramme 2-D peut être considéré comme une image  $J$  dont les dimensions spatiales, sont celles des deux axes choisis dans l'espace colorimétrique associé à  $I$ . Grâce à cette représentation, on peut analyser un histogramme 2-D en utilisant les outils d'analyse d'images en niveaux de gris [LEZ 04].



Figure 16 : Segmentation par l'histogramme 2D dans l'espace couleur  $L^*a^*. L^*b^*$

### III.3 Conclusion

La segmentation d'une image couleur, consiste à partitionner l'image en régions disjointes. En général, on suppose que les différentes couleurs présentes dans l'image, correspondent principalement aux différentes orientations et propriétés de réflexion des surfaces des objets observés ainsi qu'à la présence d'ombres [LEZ 04]. Les méthodes de segmentation, analysent ces couleurs afin de distinguer les différents objets qui composent une scène observée par une caméra couleur. Les différences entre les couleurs présentes dans l'image, sont plus ou moins bien mises en évidence selon le système de représentation de la couleur utilisé. Dans le cadre de notre étude, nous avons opté pour la morphologie mathématique suivant la ligne de partage des eaux améliorée (LPE).

## IV. Méthode de segmentation couleur adoptée

Après avoir vu les différentes méthodes de segmentation couleur, nous avons choisi après une étude bibliographique sur les techniques de segmentation en imagerie cytologique, à la segmentation par L.P.E améliorée.

### IV.1 La ligne des partage des eaux

En morphologie mathématique, l'une des principales transformations, est la ligne de partage des eaux. Celle-ci a été étendue relativement au cas des images couleurs. L'innovation principale, est venue des travaux de Meyer [Mey92] qui a proposé un algorithme dont les similarités avec les méthodes de croissance de régions, sont nombreuses. Il utilise des différences de couleur entre pixels ainsi qu'une structure de données de type file hiérarchique pour réaliser une croissance de régions rapide où les régions se propagent à partir des marqueurs. La mesure utilisée pour déterminer la probabilité d'appartenance d'un pixel à une région, et donc sa probabilité d'agrégation, s'exprime par une fonction de potentiel. Cette fonction consiste en une mesure de similarité entre un pixel et, une région voisine à laquelle il pourrait être agrégé. Les recherches qui ont été menées, se situent surtout au niveau de cette fonction de potentiel [LEZ 00].

#### IV.1.1 Fonction de potentiel couleur

La couleur des objets présents dans une image, se traduit directement par une mesure statistique de la couleur moyenne des régions et les transitions entre les couleurs des objets, s'expriment facilement par le gradient couleur. L'utilisation de la couleur suivant ces deux critères peut alors amener, à des résultats qui correspondront à des changements significatifs dans les images couleurs. Ainsi, une ligne de partage des eaux, sera précise en utilisant les deux informations suivantes : l'information locale donnée par le gradient couleur et l'information globale donnée par une mesure de la couleur moyenne des régions qui permet de traduire l'homogénéité globale de celles-ci. La fonction de potentiel est alors défini suivant la relation ci-dessous [LEZ 00]:

$$f(p, R) = (1 - \alpha) \|\overline{I_{c_1 c_2 c_3}(R)} - I_{c_1 c_2 c_3}(p)\| + \alpha \|\nabla I_{c_1 c_2 c_3}(p)\| + \alpha \|\nabla I_{c_1 c_2 c_3}(p)\|$$

Où  $\overline{I_{c_1 c_2 c_3}(R)}$  est le vecteur donnant la couleur moyenne de la région R pour l'image I dans l'espace couleur  $c_1 c_2 c_3$ ,

et Le vecteur  $I_{c_1 c_2 c_3}(p)$  donnant la couleur au point p et  $\nabla I_{c_1 c_2 c_3}(p)$  le gradient couleur au point p.

#### IV. 1.2 Méthodologie d'utilisation

L'utilisation de notre méthode de segmentation morphologique, nécessite plusieurs choix :

- Choix de l'espace couleur,
- Choix du gradient couleur,
- Choix d'une métrique pour comparer  $\overline{I_{c_1 c_2 c_3}(R)}$  et  $I_{c_1 c_2 c_3}(p)$ ,

- Choix des marqueurs,
- Choix du coefficient de corrélation  $\alpha$ .

Afin de bien réaliser ces choix, qui sont déterminants pour obtenir de bons résultats, nous définissons la méthodologie d'utilisation. Elle permet de mettre en place, une stratégie de segmentation sur des images couleurs de microscopie [LEZ 00].

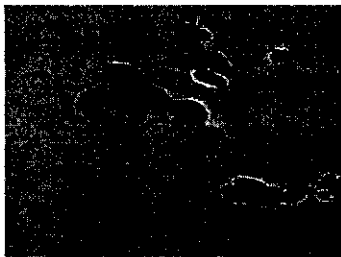
#### IV.1.3 Choix de l'espace couleur

La méthode étudiée, consiste à extraire ces objets par une segmentation manuelle (cf. Figure 17) et à déterminer quel espace couleur correspond le mieux aux transitions entre les régions (cf. Tableau 1).

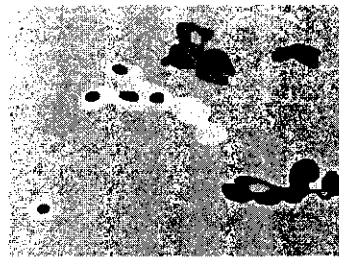
Espace	RVB	XYZ	L*u*v*	L*a*b*	HSL
Contraste	67.20	64.25	45.19	27.83	64.49

**Tableau 1** : Contraste apporté pour chaque espace couleur

Les différentes valeurs nous permettent de conclure que l'espace couleur RVB est le mieux approprié [LEZ 00].



(a) une image microscopique



(b) la région de la figure (a)

**Figure 17** : Une image de microscopie et les régions extraites manuellement.

#### IV. 1.4 Choix du gradient couleur

L'opérateur gradient utilisé, est un gradient couleur et est défini ainsi :

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

Les filtres utilisés pour calculer le gradient suivant les directions X et Y sont respectivement FX et FY :

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

#### IV. 1.5 Choix de la métrique

Le choix dépend d'informations *a priori*. La distance de Mahalanobis est mieux adaptée pour des régions qui présentent de fines variations mais pour des objets suffisamment homogènes, on utilisera la distance euclidienne. C'est donc cette dernière que nous avons choisie [LEZ 98].



#### IV.1.6 Choix des marqueurs

Pour résumer, l'extraction des marqueurs peut se faire sur une composante couleur, sur la combinaison de plusieurs composantes provenant de différents espaces couleur ou bien encore sur l'union de marqueurs. Le choix des composantes, doit être guidé par l'utilisation d'informations a priori sur les objets à extraire [LEZ 98].

- **Les marqueurs cytoplasmiques :** Les marqueurs permettent d'éliminer certains minima régionaux qui pourraient être gênants. Notre but est de trouver une composante connexe par cytoplasme et, d'imposer à la ligne de partage des eaux ces marqueurs (cf. Figure 18). Pour trouver les marqueurs, il faut étudier toutes les composantes dans plusieurs espaces couleurs (RVB, XYZ, HSL,  $L^*u^*v^*$ ,  $L^*a^*b^*$ ) [LEZ 98]. D'après le tableau 1, on a choisi l'espace RVB.

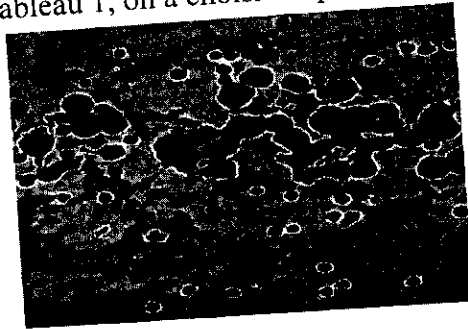


Figure 18 : un exemple de segmentation du cytoplasme des cellules

- **Le gradient et la ligne de partage des eaux pour le cytoplasme :** Toutes les techniques utilisées sont résumées dans la Figure 19.

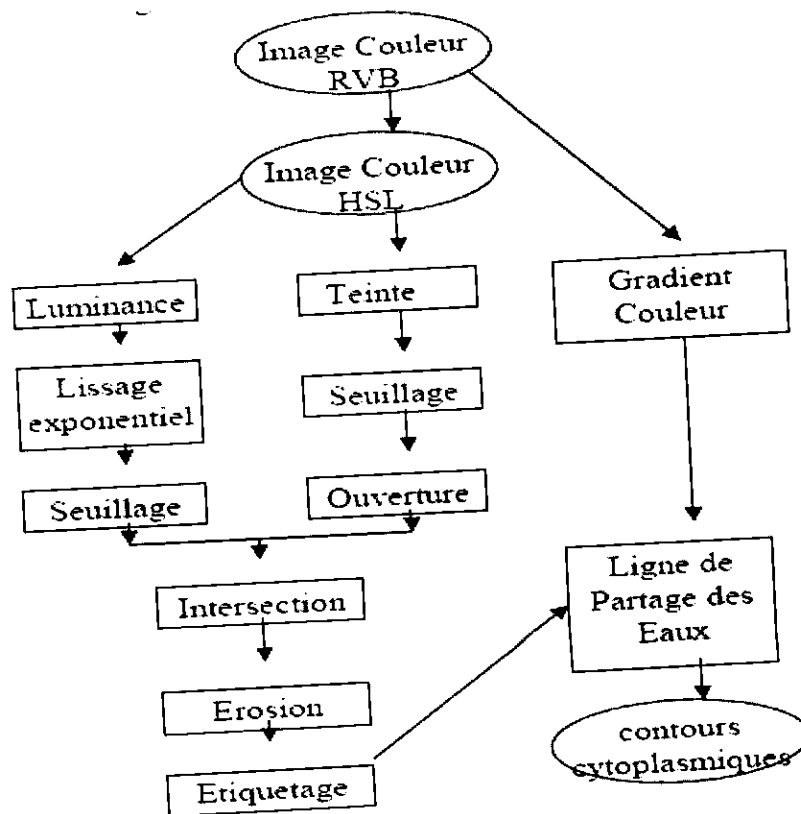


Figure 19: Synoptique de segmentation couleur du cytoplasme

- **La segmentation nucléaire :** La segmentation des noyaux est également basée, sur une méthode morphologique couleur basée sur la ligne de partage des eaux par marqueurs.
- **Les marqueurs nucléaires :** Après le choix de l'espace couleur, le seuil qui donne les meilleurs marqueurs est :  
 $T1 = (\sum f(p)) / p \text{ h et } p < S1$   
 $T2 = (\sum f(p)) / p \text{ h et } p < S2$   
 $S = S1 \text{ si } T1 < T2 \text{ et } S = S2 \text{ autrement.}$   
Avec : S1 seuil obtenu pour la variance,  
S2 seuil obtenu pour la vallée et p un point.
- **La Ligne de Partage de Eaux sur les noyaux :** Le calcul de gradient va nous permettre de voir le noyau et le cytoplasme tels qu'ils sont à l'endroit de la transition bleu-vert (noyau cytoplasme). Après étiquetage des marqueurs, on effectue la segmentation par la ligne de partage des eaux. La stratégie entière, est résumée dans la figure 22. Nous distinguons les résultats dans l'exemple des figures 20 et 21 : les noyaux des cellules isolées sont bien définis, les noyaux accolés sont détectés comme tels et les amas forment des zones plus complexes mais formant bien un seul objet.

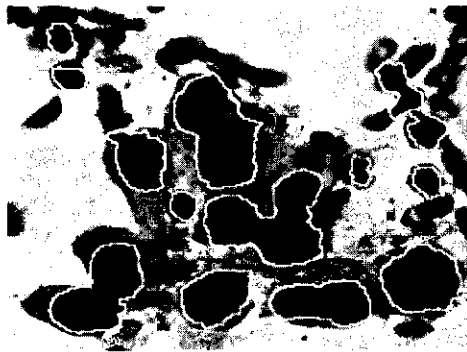


Figure 20 : Noyaux accolés



Figure 21 : Noyaux en amas

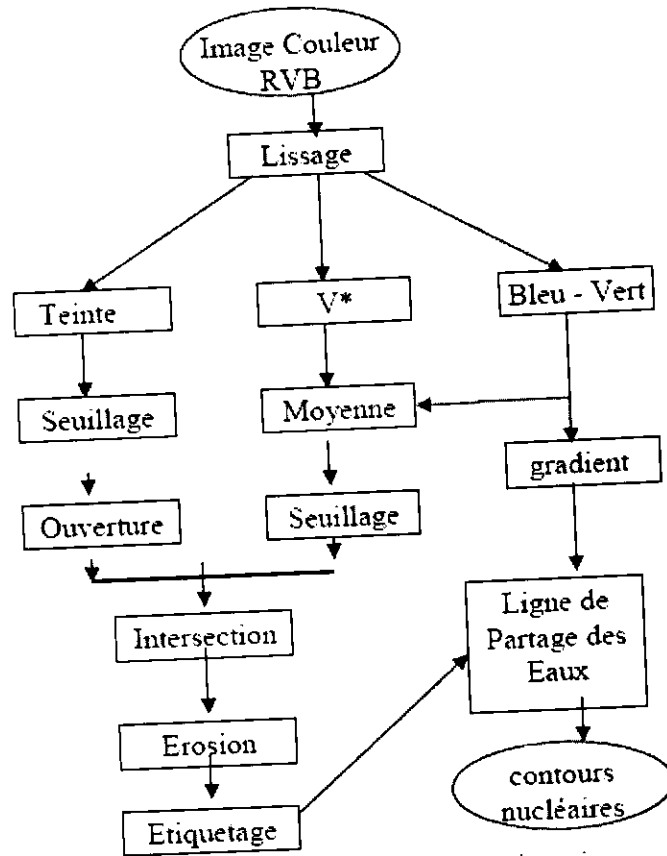


Figure 22 : Synoptique de la segmentation couleur des noyaux

- **Séparation des noyaux accolés :** Nous avons utilisé une distance de chanfrein 3x3 [LEZ 98]. Les marqueurs sont les érodés ultimes de chacun des noyaux obtenus à la précédente étape. Ces érodés sont calculés par une reconstruction par érosion de hauteur 2. Les marqueurs obtenus permettent alors de séparer plusieurs noyaux grâce à l'inverse de la fonction distance (cf. Figures 23 et 24).

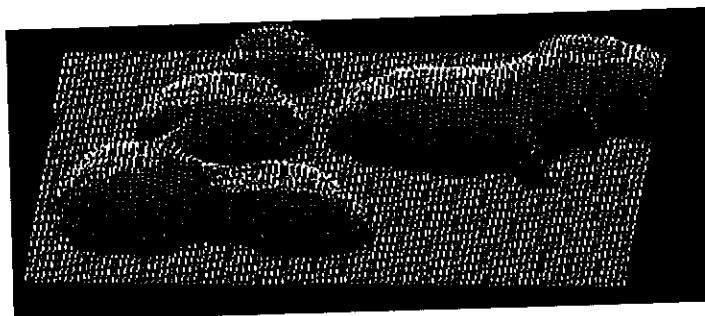


Figure 23 : Fonction distance sur des noyaux accolés en vue 3D



Figure 24: Noyaux après séparation

#### IV. 1.7 Choix du coefficient de corrélation $\alpha$

Si les contours sont très nets, on prendra le paramètre  $\alpha$  avec une valeur proche de 1. Si les contours sont moins nets et les régions peu homogènes, on prendra  $\alpha$  proche de 0.5 et si, les contours ne sont pas nets mais que les régions sont homogènes, on prendra  $\alpha$  proche de 0 [LEZ 98].

### IV. 2 Classification de pixel par analyse d'histogramme 2-D

La notion de « dimension » d'un jeu de données se réfère au nombre de descripteurs que possède chaque élément du jeu de données (pixel en imagerie 2D ou 3D). Ainsi, pour une image « traditionnelle » en niveaux de gris, la dimension est 1 (ne pas confondre avec la taille de l'image qui se réfère au nombre de plans, de colonnes et de lignes !). Pour une séquence temporelle, la dimension est le nombre (fini) de mesures effectuées au cours du temps. Pour une imagerie statique à 3 marqueurs fluorescents, la dimension est 3. Pour une image spectrale avec des spectres à 1024 canaux, la dimension est 1024 [BON 06].

#### IV. 2. 1 Espace couleur de l'histogramme

Pour réaliser un espace de couleur d'une image, plusieurs stratégies peuvent être utilisées. Les stratégies diffèrent généralement suivant la dimension des données de l'image. Une alternative intéressante à détermination du nombre de classes, et établie sur l'utilisation des histogrammes partiels (les histogrammes 2D) qui utilisent deux bandes de couleur ensemble notamment  $L^*a^*$ ,  $L^*b^*$  et  $a^*b^*$  dans  $L^*a^*b^*$  l'espace de couleur. Un autre avantage à considérer, est le fait qu'un histogramme 2D, n'est rien d'autre qu'une image de niveau de gris : les algorithmes classiques et rapides de traitement d'image de niveau de gris, peuvent donc être employés pour grouper les histogramme 2D [LEZ 04].

#### IV.2.2. Espace morphologique non supervisé

Dans cette section, l'espace morphologique non supervisé est étudié pour grouper les histogrammes 2D obtenus à partir d'images couleurs.

Un histogramme 2D, représente la projection d'un histogramme 3D sur des paires de bandes (Cf. figures 25 (a)-(b)) (les différents objets dans l'image, sont représentés dans l'histogramme autour des crêtes dominantes) [LEZ 04].

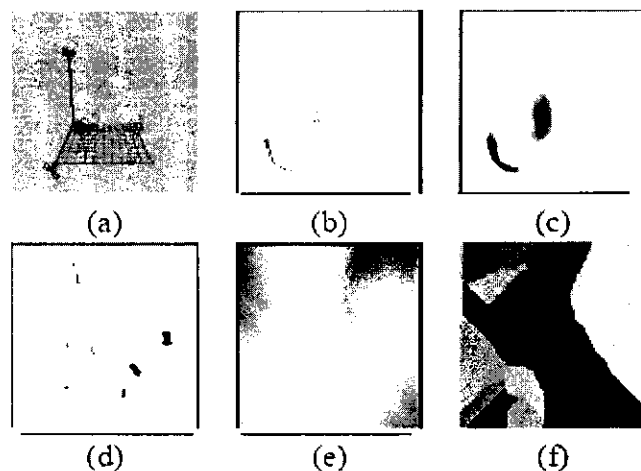


Figure 25: (a) Histogramme 3D dans l'espace couleur.  $L^*a^*b^*$

- (b) Histogramme 2D dans  $L^*a^*$
- (c) Histogramme simplifié.
- (d) les couleurs dominantes.
- (e) La fonction distance dans la couleur dominante.
- (f) Regroupement par la L.P.E.

#### IV. 2.3 Simplification d'image par lissage non linéaire d'un graphe

Une partition fine obtenue par un critère connectif quelconque est très sursegmentée [1]. Cependant, cette partition est fine car elle est très fidèle au contenu de l'image originale. Si l'on construit l'image mosaïque d'une partition fine, cela revient donc à effectuer une simplification de l'image. Effectivement, une image mosaïque d'une partition correspond à une image où chaque pixel a comme couleur la couleur moyenne de la région à laquelle il appartient. L'approche de simplification étudiée (cf. Figure.26), opère directement sur une partition. Cependant, lorsque l'on considère une partition donnée  $P_k$ , cela revient à considérer son graphe d'adjacence de régions (RAG)  $G_k$  (une hiérarchie de partitions  $H$  correspond alors à une hiérarchie de RAG) [LEZ 04].

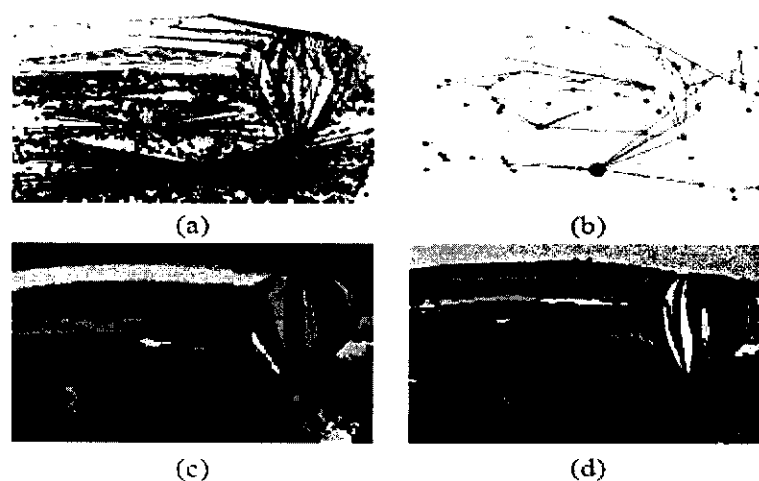


Figure 26 : Image simplifiée (a) RAG initial, (b) RAG simplifié (c) régions après la fusion, (d) filtrage morphologique.

#### IV. 2.4 Hiérarchie de partitions par fusion de régions

La méthode de simplification d'images couleurs que nous venons de présenter, est très intéressante pour obtenir une représentation simplifiée des modèles attachés à chaque région [1]. En effet, la simplification de l'image tend à rapprocher des modèles similaires et les régions devenues similaires, peuvent alors fusionner. L'idée de fusionner des régions provenant d'une partition initiale est assez ancienne, elle est à la base de beaucoup de méthodes hiérarchiques dont les pyramides irrégulières [LEZ 04]. Une stratégie de fusion de régions définie sur un RAG nécessite de définir les notions suivantes :

1) **Le modèle des régions MRI** : un modèle permet de représenter une région, mais également de définir comment est réalisée l'union de deux régions qui fusionnent.

2) **L'ordre de fusion O(Ri,Rj)** : elle associe à chaque arête du graphe, une mesure de similarité entre les noeuds. Cet ordre, est une fonction définie pour chaque couple de régions voisines et ses valeurs appartiennent à un ensemble totalement ordonné.

3) **Le critère de fusion C (Ri,Rj)** : ce critère définit si deux régions adjacentes doivent être fusionnées .

Le modèle des régions utilisé, est très simple et est constitué uniquement de la couleur moyenne de chaque région :  $\mathcal{M}_{R_i} = N_i^t$  .

L'union de deux régions devant être calculée rapidement. Elle est définie directement à partir des modèles des deux régions devant fusionner :

$$\mathcal{M}_{R_i} \cup \mathcal{M}_{R_j} = N_i^t \cup N_j^t = N_i^t + N_j^t.$$

Chaque région est décrite dans ce cas par un vecteur de quatre attributs :

$$\mathcal{M}_{R_i} = \left( \sum_{p \in R_i} C_R(p), \sum_{p \in R_i} C_V(p), \sum_{p \in R_i} C_B(p), |R_i| \right)$$

Où ; CX (p) étant la composante couleur X du pixel p. L'ordre de fusion utilise directement la similarité entre deux régions telle que nous l'avons définie pour la simplification des modèles de régions :

$$\mathcal{O}(R_i, R_j) = \mathcal{O}(N_i^t, N_j^t) = d(N_i^t, N_j^t)$$

Pour effectuer la fusion des régions répondant au critère de fusion, les régions sont classées dans une file hiérarchique prioritaire selon leur similarité établie par l'ordre de fusion, et les deux régions dont la similarité est la plus faible sont fusionnées [LEZ 04].

Nous résumons l'approche dans sa globalité par l'algorithme ci-dessous :

```

 $\lambda$  : réel;  $\lambda_{end}$  : réel;
 $\lambda \leftarrow 1$ ; Définir  $\lambda_{end}$ 
 $\mathcal{G}_\lambda = (\mathcal{N}_\lambda, \mathcal{A}_\lambda)$  pour une partition initiale  $\mathcal{P}_\lambda$ .
Tant que ( $\lambda \leq \lambda_{end}$ ) faire
    Pour les noeuds  $N_i \in \mathcal{N}_\lambda$  faire
        Simplifier le modèle du noeud  $N_i$  en ap-
        pliquant une fois le lissage non linéaire.
    Fin Pour
    Pour les arêtes  $A_l = (N_i, N_j) \in \mathcal{N}_\lambda \times \mathcal{N}_\lambda$  faire
        Si ( $\mathcal{O}(N_i, N_j) < \lambda$ ) Alors
            Ajouter  $A_l$  au noyau de contraction
             $NC_{\lambda, \lambda+1}$ 
        Fin Si
    Fin Pour
    Contracter le graphe  $\mathcal{G}_\lambda$  avec le noyau
    de contraction  $NC_{\lambda, \lambda+1}$  :  $\mathcal{G}_{\lambda+1} =$ 
     $Contraction[\mathcal{G}_\lambda, NC_{\lambda, \lambda+1}]$ 
     $\lambda \leftarrow \lambda + 1$ 
Fait
    
```

Figure 27: Algorithme de hiérarchie de partitions par simplification et fusion de régions.

#### IV.2.5 Amélioration de la L.P.E

Malheureusement la segmentation obtenue après le RAG simplifié, peut produire des bornes irrégulières des objets [LEZ 04]. Pour réduire cet effet, un filtrage morphologique est appliqué. Deux érosions successives d'étiquette sont réalisées : si un pixel a une étiquette  $j$  et n'importe lequel de ses voisins a une étiquette  $k \neq j$ , ou est non affecté, alors le pixel est classé non affecté. Des repères sont placés à tout le restant des régions après le filtrage morphologique et un algorithme de ligne de partage de couleurs est alors employé pour remplir les zones non affectées.

La figure suivante, montre les résultats de cette segmentation [LEZ 04] :

- 33 -En réponse à : Dernier enregistrement par Kamel

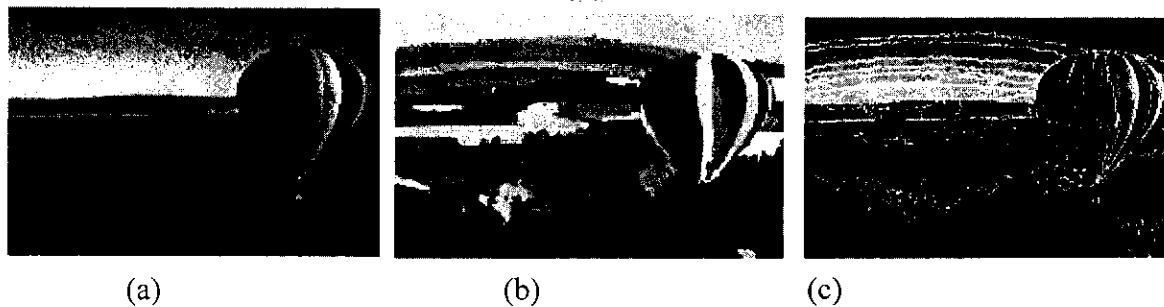


Figure 28 Segmentation après la L.P.E améliorée  
 (a) image initiale, (b) régions finales  
 (c) frontières superposées

## V. Conclusion

Ce chapitre a été consacré à l'étude de la segmentation en régions par analyse et plus en détails suivant la morphologie mathématique appliquée aux images couleurs. Nous avons présenté une méthode pour la production d'une hiérarchie de partitions en combinant un lissage sur graphe d'adjacence de régions avec une décimation de graphe. L'approche consiste tout d'abord à définir une partition initiale considérée comme fine, qui constitue la base de la hiérarchie de partitions. Cette partition fine est décimée itérativement en alternant une simplification du graphe d'adjacence par un filtrage non linéaire, opérant directement sur le graphe avec une fusion des régions similaires selon un critère de similarité défini par le niveau de la hiérarchie. L'approche étudiée, permet de réaliser de très bonnes simplifications d'images pour les niveaux inférieurs de la hiérarchie mais également d'obtenir des segmentations efficaces dans les niveaux supérieurs de la hiérarchie. De plus, aucun paramètre n'est nécessaire et seule la partition fine initiale a une influence sur la hiérarchie produite. Cette approche peut être utilisée soit dans le cadre de la simplification d'images, soit dans le cadre de la segmentation d'images, le cas de notre étude.

Le chapitre qui suit, est consacré à la caractérisation des régions extraites suivant la ligne de partage des eaux et à la classification connexionniste, en particulier suivant le modèle du perceptron.





# *CHAPITRE III*

Nous introduisons dans ce chapitre, la caractérisation morphologique des cellules, traitées suivant la morphologie mathématique ainsi que la classification suivant le modèle du perceptron.

## I. Caractérisation morphologique de la cellule

L'étape d'extraction de paramètres morphologiques est nécessaire, pour classifier nos cellules. Elle se base sur l'hyperchromatisme nucléaire, le rapport nucléo-cytoplasmique, l'anisocaryose et la difformité nucléaire.

### I.1 Description des paramètres morphologiques

Un certain nombre de critères cytologiques, permettent d'identifier la plupart des cellules cancéreuses [BEU 90]. Nous avons retenu quatre, qui figurent parmi les symptômes de cancer les plus fréquents.

- **Le rapport nucléo-cytoplasmique** : il s'agit du rapport entre la surface prise par les noyaux et celle que constitue le cytoplasme, liée à une segmentation du noyau plus importante que celle du cytoplasme. En effet, dans la plupart des cellules cancéreuses, les noyaux croissent de manière anarchique et envahissent presque toute la cellule. Cette évolution anormale des noyaux permet d'identifier la plupart des cellules cancéreuses [DJE 06].
- **L'anisocaryose** : en médecine, 'aniso' est l'inégalité, 'caryose' est la taille du noyaux, c'est-à-dire, l'irrégularité de la taille des noyaux d'une cellule à l'autre lors de la mitose (phénomène de division cellulaire).
- **La difformité nucléaire** : c'est le caractère non circulaire de la forme.
- **L'hyperchromatisme nucléaire** : il est dû à l'absence du matériel chromatinien et à la répartition de la chromatine en mottes. Cet hyperchromatisme a pour effet, sur l'image, d'une granulosité trop élevée et donc également un caractère de malignité [BEC 93].

La figure 29 représente trois images contenant trois types de cellules différentes.



(a) Noyaux isolés ou accolés

(b) Noyaux en amas

(c) Noyaux isolés granuleux

**Figure 29** : Les images des différents types des cellules (a, b, c)

La figure 30 représente quelques exemples des images à traiter dans notre application. Ces images, sont en espace couleur RVB.

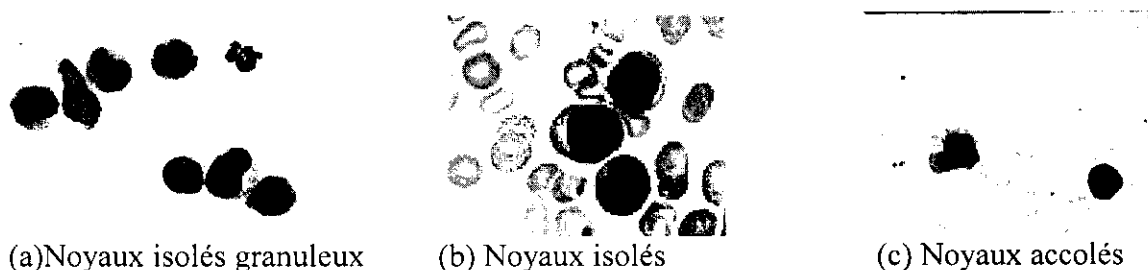


Figure 30 : Exemples d'images à traiter

## I. 2 Rapport nucléo-cytoplasmique

Comme nous l'avons dit précédemment, le rapport entre la surface occupée par les noyaux et celle occupée par le cytoplasme est un critère déterminant, dans l'analyse des cellules. A cet effet, nous avons créé, à partir de l'image prétraitée, deux images, l'une isolant les noyaux, l'autre le cytoplasme. Un simple comptage des pixels, de chacune des images suffit pour extraire le rapport nucléo-cytoplasmique.

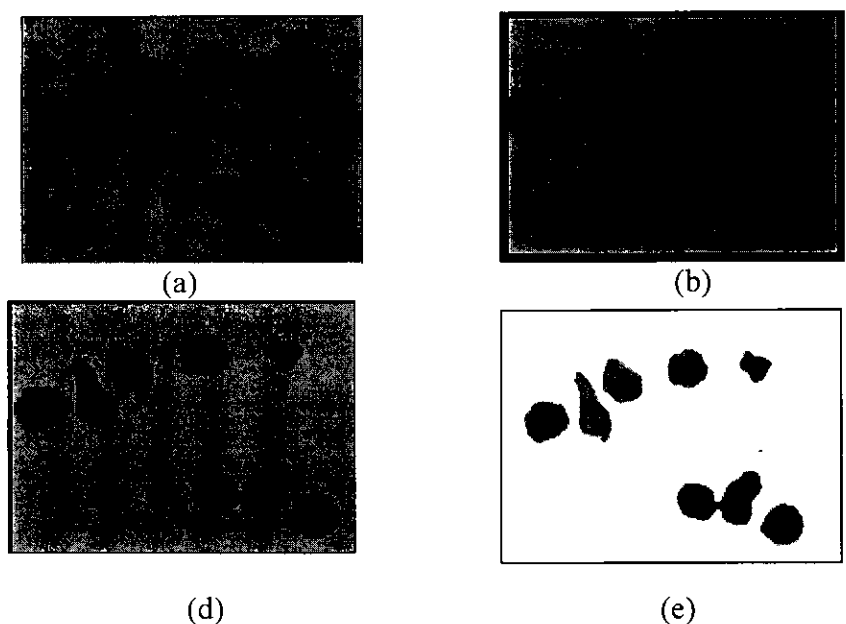
### I. 2.1 Création de l'image des noyaux

Nous avons vu que les noyaux sont des éléments foncés de l'image. L'idée première du traitement qui permettrait, de les isoler est celle d'un seuillage, éliminant les éléments clairs. Toute la démarche, consiste donc à effectuer une succession d'opérations sur l'image, permettant un seuillage aisé (où le seuil peut être choisi dans une plage assez large de valeurs de luminance). C'est cette suite organisée d'étapes que nous étudions à présent.

Nous opérons d'abord par une ouverture sur l'image, à l'aide d'un élément structurant (carré dans notre cas), la figure 31.a, représente l'image originale qui a subi une ouverture par un élément structurant carré de 7 Pixels de côté.

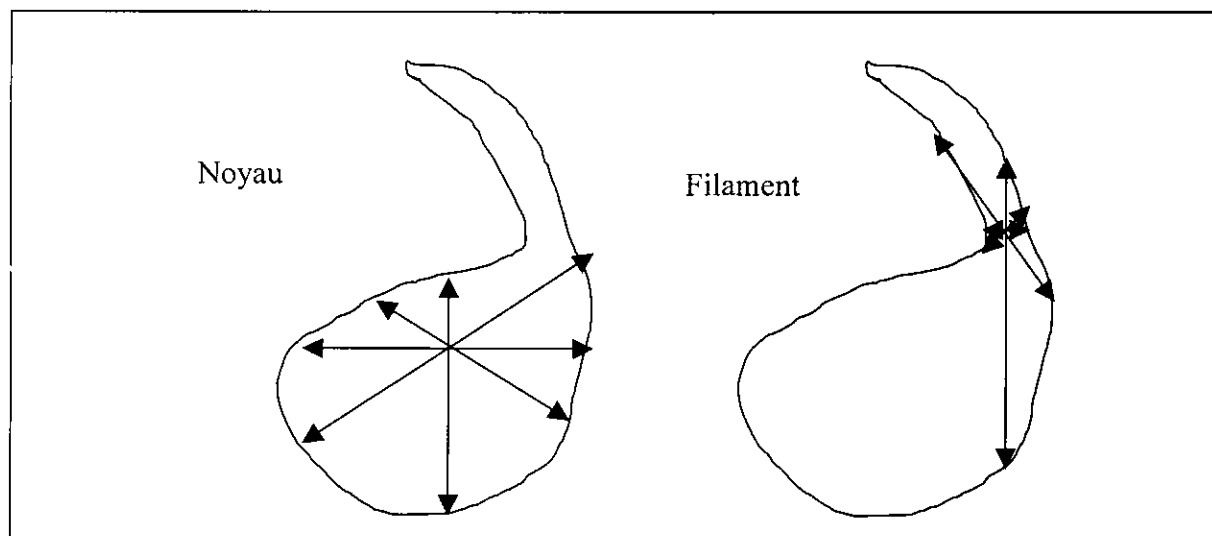
Certaines taches foncées dans des zones plus claires, ont subi le filtrage morphologique qu'effectue l'ouverture : à savoir que ces pics de luminance ont été aplanis. L'opération suivante, s'impose d'elle-même : il s'agit de la reconstruction de l'image ouverte, conditionnellement à l'image originale (cf. Figure 31.b). C'est à ce stade, que nous effectuons la principale opération du traitement qui nous intéresse : le seuillage (on ne garde que les éléments foncés de l'image), qui est donc facilité par l'éclaircissement de taches ponctuelles auparavant foncées (cf. Figure 31.c)

L'étape suivante, est de procéder à la fermeture de l'image par un élément structurant de taille de 6 pixels, remplissant ainsi les trous éventuels dans les noyaux. La figure 32 montre comment distinguer un pixel de filament d'un pixel de noyau. Pour obtenir l'image « noyaux » désirée (cf. Figure 31.e), on effectue enfin une nouvelle ouverture, par un petit élément structurant. En effet, l'élimination des filaments peut laisser des petites parties de filament non 'gommées', qui constituent alors des taches sombres, de taille trop petite pour être confondues avec un noyau.



- (a) ouverture par un carré de 7 pixels
- (d) fermeture par un carré de 6 pixels,
- (c) image seuillée et élimination des globules
- (e) image des noyaux

**Figure 31 :** Etapes permettant d'extraire l'image des noyaux de l'image originale



**Figure 32:** Critère d'appartenance à un noyau et à un filament

Les étapes permettant d'extraire l'image des noyaux sont résumés ci-dessous en considérant les valeurs typiques que nous avons utilisées :

- **Ouverture par un grand élément structurant**
  - ✓ filtrage morphologique,
  - ✓ élimination de taches isolées,
  - ✓ altération de la forme.
- **Reconstruction de l'image ouverte**
  - ✓ récupération de la forme des éléments restants.
- **Seuillage, extraction des éléments fonces (luminance<140)**
- **Fermeture par un petit élément structurant (5-6 pixels)**
  - ✓ Remplissage des trous éventuels apparus au sein des noyaux.
- **Élimination des filaments et globules ( $\alpha=15$  pixels)**
- **Ouverture par un petit élément structurant**
  - ✓ Élimination d'éventuelles taches de petite taille.

### I.2.2 Création de l'image des cytoplasmes

Nous analysons dans ce cas (cf. Figure 33), les étapes permettant d'obtenir une image ne gardant que le cytoplasme de l'image originale. On effectue d'abord une différence d'image, à savoir que l'on efface de l'image originale (cf. Figure 30.b) l'image 'noyau' obtenue par la méthode vue précédemment. Toutefois, ce qui reste n'est pas forcément le cytoplasme : on y retrouve encore certains globules rouges isolés, d'éventuels noyaux non détectés par la méthode précédente, ainsi que des filaments. On procède alors, à un seuillage ne gardant que les éléments suffisamment clairs. D'un autre côté, certaines zones très claires de l'image, proviennent de taches lors de l'acquisition de l'image, ou de légères traces sur la coupe, ne constituant pas du cytoplasme : ces zones sont également éliminées par un seuillage [BEC 93].

Pour les mêmes raisons que dans le cas des noyaux, on effectue une fermeture par un petit élément structurant, pour combler les vides, afin de bien calculer toute la surface du cytoplasme.

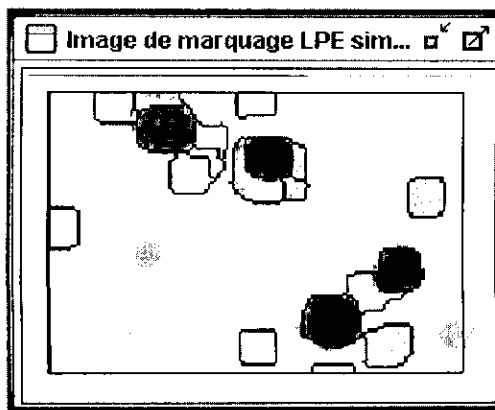


Figure 33: Image des cytoplasmes.

Les étapes permettant d'obtenir une image ne gardant que le cytoplasme de l'image originale se résument comme suit:

- **Élimination des noyaux** (dans l'image originale par la différence entre l'image originale et l'image du noyau).
- **Seuillage, élimination des éléments foncés** (luminance < 200)
  - ✓ élimination des taches foncées,
  - ✓ élimination des filaments et globules rouges,
  - ✓ élimination d'éventuels noyaux non détectés précédemment.
- **Seuillage, élimination d'éléments clairs** (luminance > 250)
  - ✓ élimination des zones non cytoplasmique, trop claires.
- **Ouverture par un petit élément structurant** (5 pixels)
  - ✓ remplissage de trous au sein des cytoplasmes.

### I. 3. Anisocaryose

L'image « noyaux » obtenue (cf. Figure 31.e) va nous servir de base pour l'étude de caractéristiques. L'anisocaryose de cette image, sera considérée comme une image binaire, puisque les valeurs de luminance dans les noyaux ne nous intéressent pas : nous distinguons simplement le blanc et le 'non blanc'.

#### I.3.1 Repérage de noyaux

Les noyaux de l'image ne sont que rarement entièrement dissociés les uns des autres; ce qui pose un problème pour les analyser un par un. Il nous a fallu implémenter un algorithme basé sur des opérations de la morphologie mathématique, afin de détecter les différents noyaux. Cet algorithme part de l'image 'noyau' et répertorie chaque noyau repéré via les coordonnées de son centre de gravité, il est donc décrit ci-dessous [BEC 93] :

- Image I = image des noyaux
  - Tant que (image I ≠ image vide)
    - image 2 = érosion (image 1)
    - quelque soit la région I incluse dans l'image 1 :
      - s'il existe une région i incluse dans l'image 2 telle que la région j incluse dans la région i,
      - alors ne rien stocker,
      - sinon, stocker les coordonnées du centre de gravité de la région i
    - image 1 = image 2

**Figure 34 :** Algorithme de calcul centre de gravité

### I.3.2 Calculs intermédiaires sur chaque noyau repéré

Une fois les noyaux repérés et stockés via les coordonnées de leur centre de gravité, on calcule, pour chaque centre de gravité, un certain nombre de distances caractéristiques, de la même manière que nous avons décrite dans la partie 'création de l'image 'noyaux'' pour la suppression des filaments. Il s'agit donc, de calculer la distance du centre de gravité aux bords de sa région, dans un certain nombre de directions. La figure suivante (cf. Figure 35) montre le risque de considérer un noyau comme fort allongé dans une direction (direction x dans la figure 34), alors qu'il jouxte simplement un autre [BEC 93].

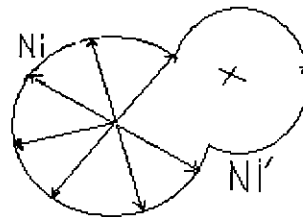
La taille moyenne d'un noyau peut être caractérisée par la moyenne suivante :

$$\text{Taille moyenne du noyau } i = t_{mi} = \text{moy}(d_{ij})$$

Où  $D_{ij}$  : est la distance caractéristique du noyau  $i$  dans la direction  $j$ .

La variance des distances caractéristiques, sur toutes les directions significatives, peut traduire le caractère difforme du noyau :

$$\text{Diformité du noyau } i = d_{fi} = \text{var}(d_{ij})$$



**Figure 35:** Distance caractéristique  $x$  non significative dans l'image de deux noyaux accolés  $N_i$  et  $N_i'$ .

Nous avons choisi, pour quantifier le caractère irrégulier de la taille des noyaux, de considérer la variance, sur tous les noyaux de l'image, de leur taille moyenne  $t_{mi}$ . De plus, afin de ne pas privilégier des images prises avec des grossissements élevés, ou des tissus à gros noyaux, cette variance est ramenée à une valeur réduite, via une division par le carré de la taille moyenne  $t_m$  de tous les noyaux de l'image [BEC 93]. Plus formellement :

$$\text{Taille moyenne des noyaux de l'image} = t_m = \text{moy}(t_{mi})$$

$$\text{Anisocaryose} = \text{var}(t_{mi})/t_m^2$$

#### I. 4. Difformité nucléaire

Après avoir effectué les étapes précédentes (repérage des noyaux et calculs intermédiaires sur chaque noyau repéré), on peut représenter le paramètre 'difformité des noyaux' par la moyenne, sur tout les noyaux, de la difformité dfi du noyau i. Ce paramètre est aussi ramené à une valeur, comme le montre la formule [BEC 93] :

:

$$\text{Difformité nucléaire} = \text{moy}(\text{dfi}) / \text{tm}^2$$

#### I.5 Hyperchromatisme nucléaire

Rappelons que le cancer provoque un accroissement anormal de l'activité nucléaire (cf. Figure 36), ce qui se traduit à l'image par des noyaux 'granuleux', présentant une texture particulière [BEC 93]. Il s'agit donc de travailler à nouveau à partir de l'image 'noyaux' obtenue en I. 2 .1 (création de l'image de noyaux). Maintenant, les valeurs de la luminance originales sont à prendre en compte, puisque le but est de quantifier les sauts de luminance au sein des noyaux.

Les étapes suivantes du processus, sont des opérations bien connues, utilisant la morphologie mathématique, dont le but est de ne garder que le caractère irrégulier de la luminance des noyaux. Ces étapes sont :

1. obtention de l'image 'noyau' originale,
2. fermeture par un petit élément structurant (5 pixels),
3. différence entre les 2 images précédentes,  
✓ obtention des creux de luminance
4. calcul du paramètre 'hyperchromatisme nucléaire'

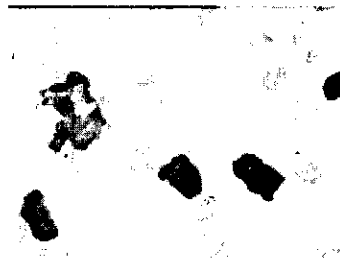


Figure 36 : Existence de l'hyperchromatisme nucléaire dans les cellules ci-dessus

A la fin, Ces quatre paramètres nous permettent de bien classifier un tissu cancéreux. On peut remarquer aussi, que plus ces valeurs seront élevées, plus les paramètres indiqueront une malignité cellulaire.

## II. Classification connexionniste

La deuxième partie de ce chapitre, présente le fonctionnement générale d'un réseaux de neurones ainsi que quelques réseaux célèbres, afin d'englober notre connaissance sur la classification connexionniste [4].

### II. 1 Qu'est-ce qu'un réseau de neurones ?



Tout d'abord, ce que l'on désigne habituellement par réseau de neurones, est en fait un réseau de neurones artificiels basé sur un modèle simplifié de neurone [5]. Ce modèle permet certaines fonctions du cerveau, comme la mémorisation associative, l'apprentissage par l'exemple et le travail en parallèle. Cependant, le neurone artificiel est loin de posséder toutes les capacités du neurone biologique (cf. Figure 37) [YZI 05]. Les réseaux de neurones biologiques sont ainsi, beaucoup plus compliqués que les modèles mathématiques et informatiques.

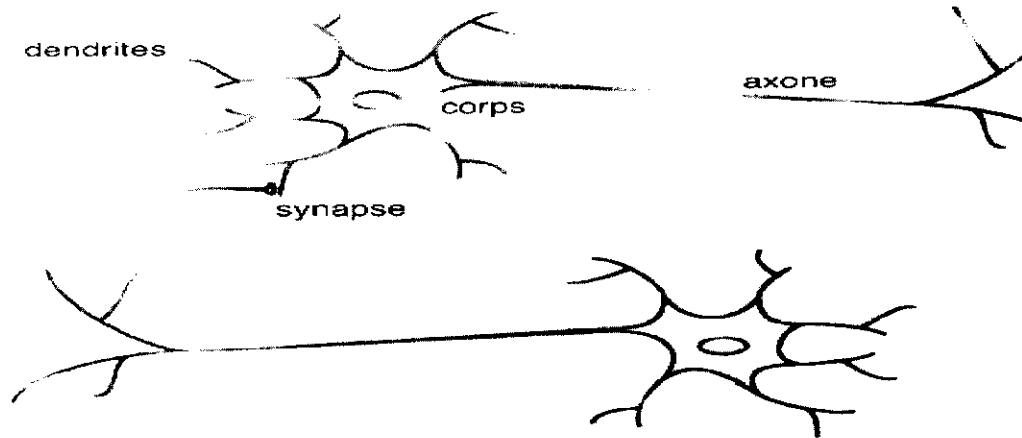


Figure 37 : Schéma d'un neurone biologique

Il n'y a pas de définition universelle du réseau de neurones [5]. On considère généralement qu'un réseau de neurones est constitué, d'un grand ensemble d'unités (ou neurones), ayant chacune une petite mémoire locale. Ces unités sont reliées par des canaux de communication (les connexions, aussi appelées **synapses** d'après le terme biologique correspondant.), qui transportent des données numériques. Les unités peuvent uniquement agir sur leurs données locales et sur les entrées qu'elles reçoivent par leurs connexions (cf. Figure 38).

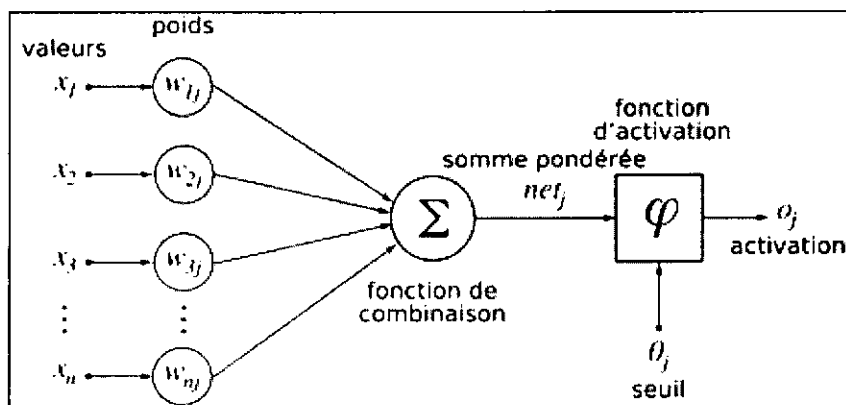


Figure 38 : Structure d'un neurone artificiel

## II.2 Modélisation d'un neurone

Si on prend un neurone à trois connexions en entrée, le reliant à trois autres neurones; il reçoit de l'information provenant de chacun de ces trois neurones. Les valeurs qu'il reçoit ainsi en

entrée, par chacune de ses connexions sont respectivement notées  $e_1$ ,  $e_2$  et  $e_3$  : ce sont les **entrées** du neurone.

Toutes les connexions n'ont pas une importance égale pour le neurone. Certaines sont plus importantes que d'autres. Le poids  $w$  affecté à chaque connexion, mesure cette importance relative : le **poids** est proportionnel à l'importance de la connexion. En fait, tout se passe comme si le neurone ne recevait qu'une entrée  $E$  et que celle-ci prenait la valeur  $E = w_1 * e_1 + w_2 * e_2 + w_3 * e_3$ . Une fois l'entrée connue, le neurone effectue une opération qui dépend de  $E$ . Cela revient à dire, qu'il applique une fonction  $f$  à la valeur  $E$ . Cette fonction  $f$  est appelée **fonction d'activation** (cf. figure 38). Le choix de cette fonction  $f$ , se révèle être un élément constitutif important des réseaux de neurones. Ainsi, l'identité est rarement suffisante et des fonctions non linéaires et plus évoluées sont nécessaires [5].

### II.3 Connectivité

La connectivité des réseaux est la manière dont les neurones, sont connectés entre eux. Elle peut être totale (tous les neurones sont connectés entre eux) ou organisée par couche (les neurones d'une couche ne sont connectés qu'à la couche précédente en entrée et à la couche suivante en sortie). Il existe des réseaux **monocouches** ou **multicouches** [6].

### II.4 Apprentissage supervisé / non supervisé

Une caractéristique des réseaux de neurones, est leur capacité à apprendre. Mais cette connaissance n'est pas acquise dès le départ. La plupart des réseaux de neurones apprennent par l'exemple en suivant un algorithme d'apprentissage. Il y a deux algorithmes principaux : lors d'un apprentissage supervisé, les résultats corrects (c'est-à-dire les valeurs que l'on désire que le réseau obtienne en sortie) sont fournis au réseau, si bien que celui-ci peut ajuster ses poids de connexions pour les obtenir. Après l'apprentissage, le réseau est testé en lui donnant seulement les valeurs d'entrée mais pas les sorties désirées, et en regardant si le résultat obtenu, est proche du résultat désiré. Lors d'un apprentissage non supervisé, on ne fournit pas au réseau les sorties que l'on désire obtenir. On le laisse évoluer librement jusqu'à ce qu'il se stabilise [DRE 02].

### II.5 Calcul des poids synaptiques

La **rétro propagation** (cf. Figure 39), est une méthode de calcul des poids (aussi appelés poids synaptiques du nom des **synapses**, terme désignant la connexion biologique entre deux neurones) pour un réseau à apprentissage supervisé qui consiste à minimiser l'erreur quadratique de sortie (somme des carrés de l'erreur de chaque composante entre la sortie réelle et la sortie désirée) [DRE 02].

### II.6 Quelques réseaux célèbres

Il existe de nombreux réseaux de neurones actuellement. Personne ne sait exactement combien. De nouveaux réseaux (ou du moins des variations de réseaux plus anciens) sont inventés chaque semaine. On en présente ici, les classiques [DRE 02].

#### II.6.1 Le perceptron

C'est un des premiers réseaux de neurones, conçus en 1958 par Rosenblatt [DRE 02]. Il est linéaire et monocouche. Il est inspiré du système visuel. La première couche (d'entrée), représente la rétine. Les neurones de la couche suivante (unique, d'où le qualificatif de monocouche) sont les cellules d'association, et la couche finale représente les cellules de décision. Les sorties des neurones ne peuvent prendre que deux états (-1 et 1 ou 0 et 1). Seuls les poids des liaisons entre la couche d'association et la couche finale, peuvent être modifiés. La règle de modification des poids utilisée est la règle de Widrow-Hoff [DRE 02] :

Si la sortie du réseau (donc celle d'une cellule de décision) est égale à la sortie désirée, le poids de la connexion entre ce neurone et le neurone d'association qui lui est connecté n'est pas modifié. Dans le cas contraire, le poids est modifié proportionnellement. À la différence entre la sortie obtenue et la sortie désirée suivant la relation ci-dessous :

$$w \leq w + k (d - s)$$

Où  $s$  est la sortie obtenue,  $d$  la sortie désirée et  $k$  une constante positive. En 1969, Papert et Minsky [DRE 02] ont démontré les limites du perceptron classique, incapable, par exemple de simuler la fonction ou exclusif (xor).

### II.6.2 Les perceptrons multicouches (PMC)

Ils sont une amélioration du perceptron comprenant une ou plusieurs couches intermédiaires dites couches cachées, dans le sens où elles n'ont qu'une utilité intrinsèque pour le réseau de neurones et pas de contact direct avec l'extérieur. Chaque neurone n'est relié qu'aux neurones des couches directement précédente et suivante, mais à tous les neurones de ces couches (cf. Figure 39).

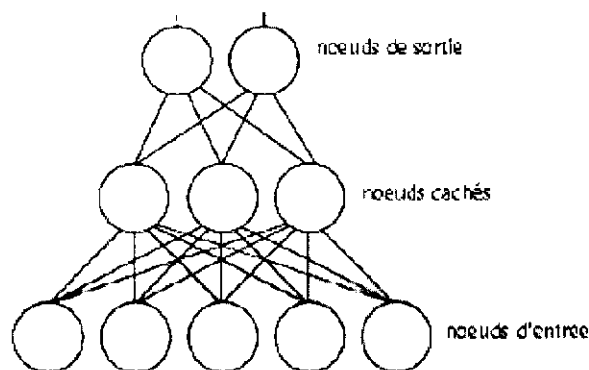


Figure 39 : Schéma représentant les PMC

Les PMC utilisent, pour modifier leurs poids, un algorithme de rétropropagation du gradient qui est une généralisation de la règle de Widrow-Hoff. Il s'agit toujours de minimiser l'erreur quadratique. On propage la modification des poids de la couche de sortie, jusqu'à la couche d'entrée. Les PMC agissent comme un séparateur non linéaire et peuvent être utilisés pour la classification, le traitement de l'image (dans notre cas c'est l'image cytologique) ou l'aide à la décision [6].

### II. 6.3. Les réseaux de Hopfield

Il s'agit d'un réseau constitué de neurones à deux états (-1 et 1, ou 0 et 1), dont la loi d'apprentissage est la règle de Hebb (1949) [DRE 02], qui veut qu'une synapse améliore son activité si et seulement si l'activité de ses deux neurones est corrélée (c'est-à-dire que le poids d'une connexion entre deux neurones augmente quand les deux neurones sont activés au même temps).

### II. 6.4 Les réseaux de Kohonen

Contrairement aux réseaux de Hopfield où les neurones sont modélisés de la façon la plus simple possible, on recherche ici un modèle de neurone plus proche de la réalité. Ces réseaux sont inspirés des observations biologiques du fonctionnement des systèmes nerveux de perception des mammifères [DRE 02].

Une loi de Hebb modifiée (tenant compte de l'oubli) est utilisée pour l'apprentissage. La connexion est renforcée dans le cas où les neurones reliés ont une activité simultanée, et diminuée dans le cas contraire (alors qu'il ne se passait précédemment rien dans ce cas). Ceci se résume par la formule suivante [DRE 02] :

$$d w / dt = k S e - B(S) w$$

### II.7 Le modèle perceptron

Il s'agit d'un réseau de neurones inspiré du système visuel possédant deux couches de neurones : une première couche relative à la perception et une seconde liée à la prise de décision. C'est le premier système artificiel capable d'apprendre par expérience.

L'objectif initial, celui poursuivi par le cognitivisme, est l'analyse des différentes étapes du processus en jeu dans la production d'une pensée rationnelle afin de tenter d'en reproduire les aspects les plus complexes. La difficulté majeure, repose cependant sur l'impossibilité de bien discerner les différentes étapes d'un tel processus. Malgré des résultats probants, force est de reconnaître que la pensée produite reste bien pâle devant celle de l'homme, indéfectible constat de l'intelligence artificielle [6].

#### II.7.1 Notion de perceptron

##### II.7.1.1 Le modèle

A l'image du neurone biologique (cf. figure 37), véritable support de notre perception des sens, le perceptron se révèle être le support de la perception des informations. On peut ainsi définir un perceptron linéaire à seuil possédant un nombre quelconque d'entrées et une sortie.

Une entrée est un booléen qui indique s'il y a une impulsion électrique et est associée à un poids qui représente l'amplitude de l'impulsion. La sortie est égale à 1 si la somme des poids est strictement supérieure à une valeur seuil. Les entrées représentent donc les dendrites, la sortie l'axone et le perceptron le neurone [YZI 05].

##### II.7.1.2 L'utilisation

L'algorithme le plus simple utilisant un perceptron, est l'algorithme par correction d'erreurs. Il s'agit de présenter au perceptron différents vecteurs d'entrées associées à un booléen. On appelle vecteur d'entrée, un vecteur qui initialise chaque entrée du perceptron à 0 ou 1. C'est l'ensemble des vecteurs d'entrées qu'on appelle échantillon. Le booléen représente la classe associée au vecteur. Tous les poids sont initialisés à une valeur aléatoire; cependant, au moins un poids doit avoir une valeur non nulle pour que la sortie ne soit pas toujours égale à zéro [YZI 05].

L'algorithme consiste, à présenter au perceptron un certain nombre de fois ces vecteurs d'entrée. Lorsqu'un vecteur est présenté au perceptron, les poids sont modifiés à chaque fois que la sortie est différente de la valeur de la classe. L'apprentissage est terminé une fois que tous les vecteurs ont été présentés sans modification des poids.

Il convient de remarquer qu'il s'agit effectivement d'un algorithme par correction d'erreurs car les poids ne sont modifiés que si la sortie calculée est différente de la sortie attendue. La règle de modification des poids est la suivante :

$$W_i <= W_i - (C - O) \cdot X_i$$

$W_i$  : représente le poids de l'entrée  $i$   
 $C$  : la classe  
 $O$  : la sortie calculée  
 $X_i$  : l'entrée  $i$

### II.7.1.3 Les limites de ce modèle

Ce modèle a rapidement montré ces limites. En effet, l'algorithme exige un échantillon linéairement séparable, ce qui est rarement le cas des données que l'on veut faire apprendre artificiellement. Dans le cas où l'échantillon n'est pas linéairement séparable, l'algorithme ne se terminera pas. C'est pourquoi, les réseaux de neurones multicouches ont été mis en place. Cette contrainte ne permettra donc pas l'apprentissage du XOR, ou exclusif, par cet algorithme basé sur la notion de perceptron. Il apparaît évidemment impossible de séparer linéairement le XOR en 2 classes distinctes (cf. Figure.40) [YZI 05].

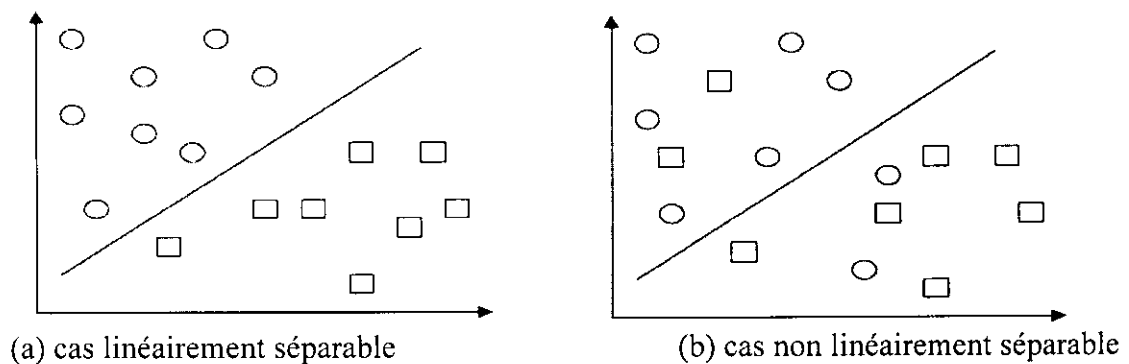


Figure 40 : Deux cas de séparation

Pour améliorer l'apprentissage du perceptron, Widrow et Hoff ont proposé une nouvelle règle d'apprentissage plus proche des mécanismes d'optimisation [YZI 05]. Il s'agit non seulement de trouver la frontière séparant les classes lorsqu'elle existe, mais aussi de trouver la meilleure solution qui minimise l'erreur lorsque cette séparation linéaire n'est pas possible.

## II.8 Classification par le réseau PMC

Le réseau de neurones ou perceptron multicouche PMC est un ensemble de neurones interconnectés. Cela constitue une extension du modèle de perceptron. Dans le réseau multicouche, les neurones sont disposés en couches successives. Dans notre cas, nous n'étudierons que les réseaux à 3 couches (cf. figure 41) :

- Une couche d'entrée représentant les diverses impulsions électriques que reçoit le réseau,
- une couche cachée qui définit l'ensemble des neurones intermédiaires,
- la couche de sortie qui définit la réponse des neurones par rapport aux impulsions d'entrée.

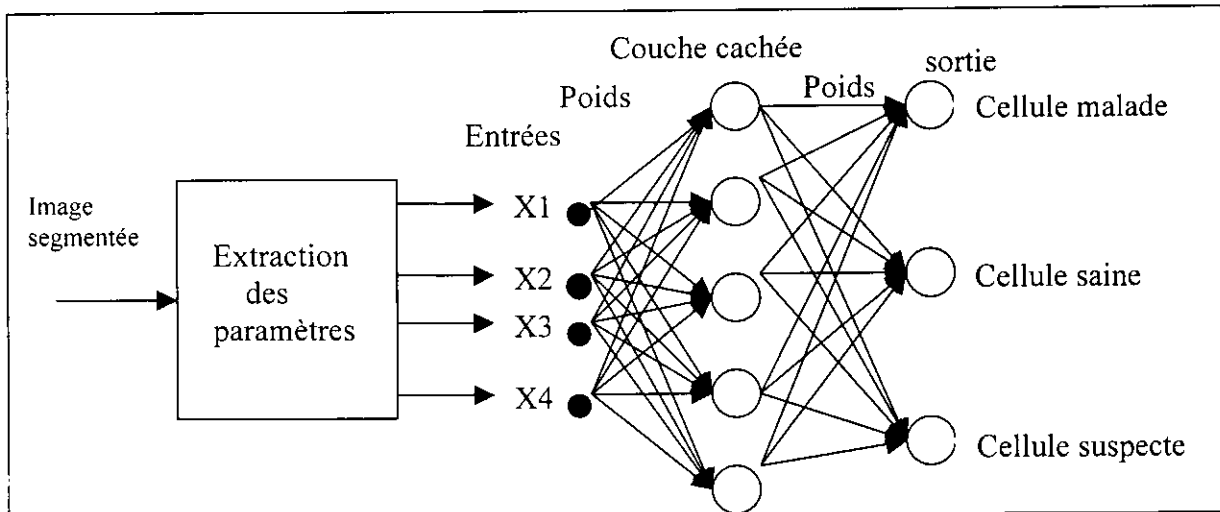


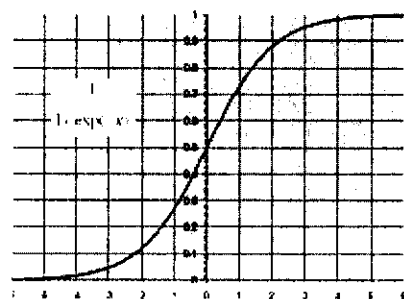
Figure 41 : Schéma expliquant le déroulement de réseau PMC

**II.8.1Création de la base d'apprentissage**

Lors de la phase d'apprentissage, (cf. Figure 43) le neurone modifie ses poids. Nous allons maintenant expliquer comment cette modification s'effectue en amont du neurone, par la règle de rétropropagation du gradient et la règle de Hebb. L'algorithme de rétropropagation du gradient, consiste à déterminer l'erreur commise par chaque neurone puis à modifier la valeur des poids pour minimiser cette erreur. Il faut effectuer des rétropropagations jusqu'à ce que l'erreur quadratique moyenne devienne inférieure à un certain seuil. L'erreur quadratique est définie comme la somme sur les sorties des carrés des différences entre l'entrée obtenue et l'entrée désirée soit :

$$E = \sum (S_i - Y_i)^2$$

Le réseau est représenté par 3 couches de neurones, chaque neurone d'une couche étant relié à tous les neurones de la couche suivante par une liaison pondérée. Chaque neurone intègre les données grâce à une fonction non linéaire qui permet l'apprentissage de classes non linéairement séparables. On utilise ici la fonction sigmoïde suivante :



$$s(x) = \frac{1}{1 + e^{-x}}$$

$$ds(x) = s(x)(1 - s(x))$$

Figure 42 : Représentation de fonction sigmoïde

On effectue tout d'abord une propagation de l'entrée à travers le réseau. Considérons S le vecteur de sortie obtenu après la propagation de l'entrée X à travers le réseau, et Y le vecteur des sorties que l'on aurait voulu obtenir. Pour la couche de sortie considérons le neurone i ; on obtient donc l'erreur pour ce neurone : S[i]-Y[i].

On appelle  $p$  la taille de la couche cachée. Le neurone  $i$  a alors, reçu les  $p$  sorties des neurones de la couche cachée modulées de leur poids respectifs. Ce sont ces poids qu'il nous faut modifier. Pour relativiser l'importance d'un poids par rapport à un autre, dans l'erreur totale il faut moduler sa modification par la sortie du neurone auquel il est rattaché [YZI 05]. Il faut maintenant utiliser le gradient de l'erreur afin de pouvoir à terme faire converger la valeur du poids. On obtient alors la formule suivante :

La règle de modification des poids pour un exemple  $k$  :

$$W_{ij}(k) = W_{ij}(k-1) - \text{pas} \times D_i \times O_j \times ds(I_i)$$

$$D_i = (S_i - Y_i) \text{ pour la couche de sortie}$$

Avec :

$W_{ij}$  : le poids entre le neurone  $i$  d'une couche et le neurone  $j$  de la couche précédente

pas : pas de l'apprentissage

$O_j$  : sortie du neurone  $j$

$I_i$  : entrée du neurone  $i$ :  $I_i = \sum_j W_{ij} \cdot O_j$

$S_i$  : sortie obtenue

$Y_i$  : sortie désirée

$ds$  : fonction sigmoïde dérivée

Pour la couche cachée, déterminer l'erreur commise par chaque neurone n'est pas instantanée ; puisque l'on ne dispose pas de la valeur théorique qu'il devrait y avoir. La méthode, consiste à remarquer que l'erreur commise par un neurone de la couche cachée est la somme des erreurs des neurones de la couche de sortie modulée par les poids respectifs des liaisons entre les deux couches. Ainsi l'erreur pour le neurone  $i$ , de la couche cachée devient [YZI 05]:

$$\sum_h W_{hi} \cdot (S[h] - Y[h]) \quad h \text{ appartenant à la couche de sortie.}$$

On obtient alors la formule de modification des poids suivante :

$$W_{ij}(k) = W_{ij}(k-1) - \text{pas} \times D_i \times O_j \times ds(I_i)$$

$$D_i = \sum_h D_h \cdot W_{hi} \text{ pour la couche cachée}$$

Après avoir présenté les étapes de l'algorithme de rétropropagation du gradient, nous résumons alors ces étapes sous un algorithme plus précis (cf. Figure.43) :

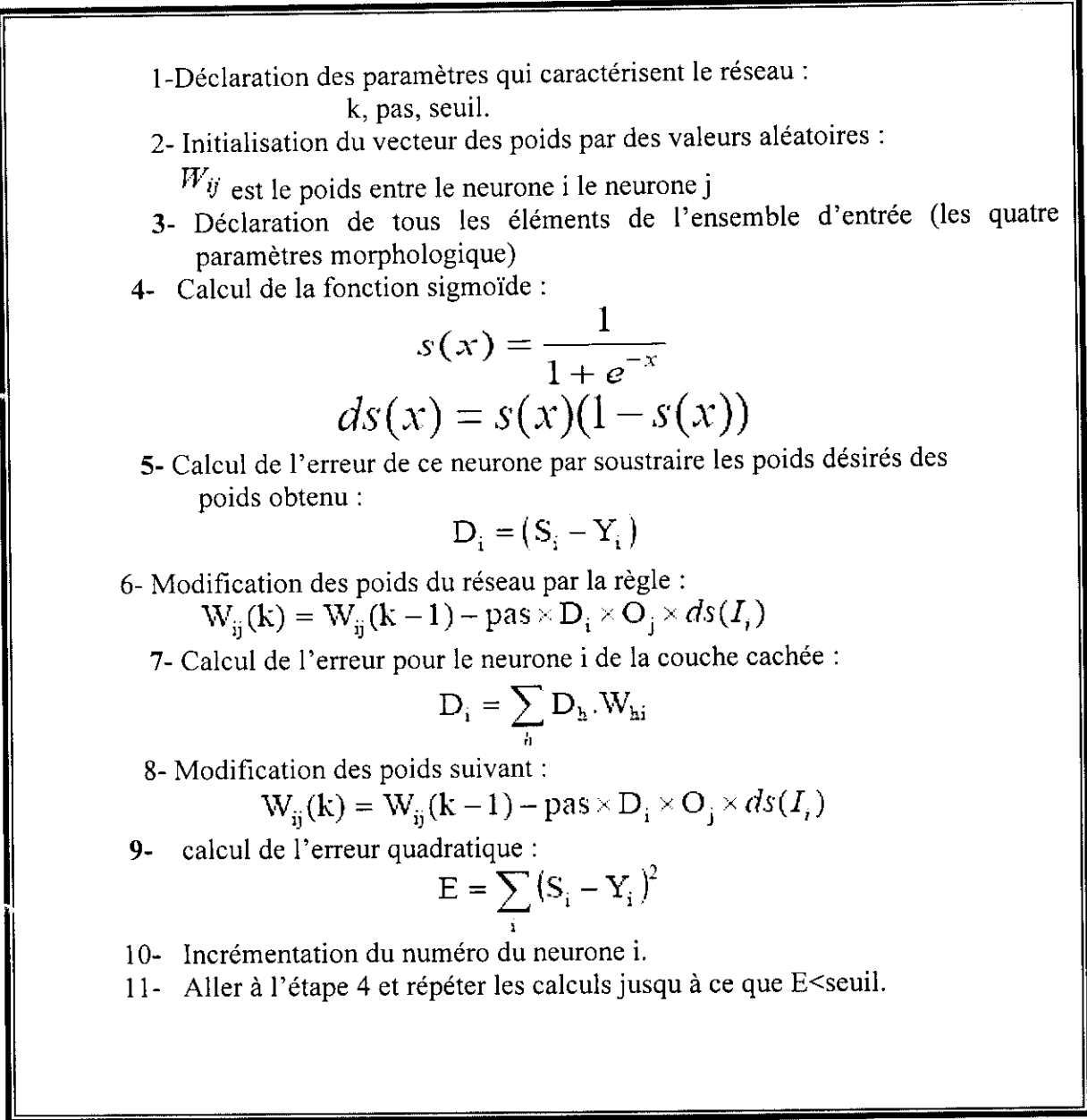


Figure 43: Algorithme d'apprentissage de réseau PMC

Nous résumons le principe de classification par la méthode connexionniste PMC dans le schéma suivant :



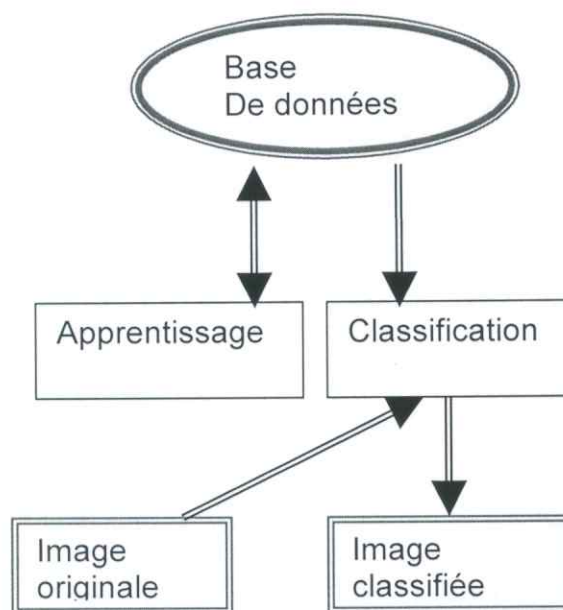


Figure 44 : principe de classification par le PMC



### II.8.2. Le réseau de neurone implanté

Les neurones sont regroupés au sein d'une classe **couche** composée de tous les neurones d'une couche. La classe **réseau**, composée de trois couches représente le réseau de neurones dans son ensemble et possède deux méthodes :

La méthode **calculSortie** qui, à partir d'un tableau de valeurs d'entrées, parcourt le réseau de neurones et détermine la sortie associée à cette entrée. Cette sortie est retournée par le biais d'un tableau passé en argument [YZI 05].

La méthode **apprendre** effectue l'apprentissage de l'exemple passé en paramètres, en fonction de la sortie désirée et du pas d'apprentissage également passé en paramètre. Cette méthode effectue un nombre suffisant de rétropropagations, permettant à l'erreur quadratique d'être inférieure au seuil voulu.

Le programme principal, comporte une méthode **rétropropagation** qui prend comme paramètres le nombre d'itération, le pas de l'apprentissage, le réseau et le fichier contenant les exemples. Au delà de la création d'un réseau de neurones, il faut effectuer différentes opérations qui peuvent rendre le réseau plus ou moins performant. En effet, il faut par exemple initialiser les poids de façon à ne pas influencer le réseau dans une direction ou une autre.

### **III. Conclusion**

Pour longtemps l'informatique est apparue comme une science purement formelle où les seules règles de la logique permettaient au programme de fonctionner indépendamment des données de l'environnement extérieur.

A contrario, aujourd'hui, il n'est plus de domaines qui ne soient totalement étrangers à l'outil informatique (domaine médical, industriel, aéronautique, télédétection...). Ils s'appuient tous sur les richesses de cet outil technologique.

A cet égard, les réseaux de neurones se trouvent être une application concrète d'interaction entre le programme informatique et un milieu : il s'agit d'apprendre et de reconnaître les dans le cas de notre projet, les cellules pathologiques, suspectes ou saines. Parti d'un modèle biologique bien connu, celui des neurones, nous en avons étudié l'adaptation informatique : le neurone formel ; le choix d'un programme de classification cytologique étant pour nous l'occasion d'étudier les relations entre un tel système et un environnement numérique.

Le chapitre suivant, présente notre environnement de travail, les traitements réalisés sur des images cytologiques prélevées sur des cas de leucémie ainsi que des discussions sur les résultats aboutis.



# *CHAPITRE IV*

Ce dernier chapitre, présente les étapes de la méthodologie développée précédemment. L'interprétation des résultats, s'obtient après une segmentation par la ligne de partage des eaux améliorée, une caractérisation morphologique une classification connexionniste par le modèle du perceptron (réseau de neurones) sur des images cytologiques indiquant des prélèvements leucémiques.

## I. Environnement de travail

### I.1 Matériel utilisé

Notre logiciel a été réalisé sur un PC dans les caractéristiques sont les suivantes:

- Pentium IV,
- CPU 1.70 GHz,
- RAM de 256 Mo,
- Disque dure de capacité 40 Go,
- Carte graphique S3 Graphiques.

### I.2 Description des images

Les images manipulées dans notre logiciel, sont des images couleur de format PNG. Ce format PNG (*Portable Network Graphics*, ou format *Png*) est un format de fichier graphique **bitmap** (raster) ; il permet de stocker des images en noir et blanc (jusqu'à 16 bits par pixels de profondeur de codage), en couleurs réelles (*True color*, jusqu'à 48 bits par pixels de profondeur de codage) ainsi que des images indexées, faisant usage d'une palette de 256 couleurs [7].

### I.3 Langage utilisé

Nous avons utilisé le langage orienté objet JAVA [7], développé par la société Sun. La syntaxe générale est très proche de celle du C, mais Java n'est pas une surcouche du C et la syntaxe est beaucoup plus claire que celle du C++.

### I.4 Les avantages de JAVA

Les avantages de Java son nombreux. Le byte code, assure à Java une portabilité complète vers de très nombreux systèmes. L'importance de l'api de base offre tous les services de base, notamment pour la construction des interfaces graphiques. La 3ème force de Java, c'est son adaptabilité dans de nombreux domaines, autant pour le web que pour les systèmes embarqués.

### I.5 Liste des systèmes disposant d'une implémentation de Java

Les systèmes disposant d'une implémentation de Java se résument suivant;  
Windows 3.1/95/98/NT4/2000/ME/XP/CE, MacOS, Solaris, Linux, AIX, OS/2, IRIX, UnixWare, HP/UX, Digital Unix, AmigaOS, BeOS, OpenVMS, FreeBSD, SunOS, RiscOS.

### I.6 Environnement de Java

Le langage java est utilisé suivant l'environnement de JRE, le JDK et le SDK [7] .

- **JRE** : Java Runtime Environment, c'est le kit destiné au client pour pouvoir exécuter un programme Java. Il se compose essentiellement d'une machine virtuelle Java (JVM) capable d'exécuter le byte-code et les bibliothèques standard de Java.
- **SDK** : Standard Development Kit, c'est le kit destiné au programmeur. Ce kit est composé d'un JRE, d'un compilateur et de nombreux programmes utiles, des exemples de programmes Java, les sources de toutes les classes de l'api.
- **JDK** : Java Development Kit, c'est l'ancien nom du SDK. Le JDK a été renommé SDK depuis la version 1.2.2.

## II. Présentation du logiciel cytomorph

cytomorph est un logiciel d'analyse d'image qui a pour objectif de faire une segmentation morphologique par la LPE améliorée des images couleurs cytologiques, une caractérisation morphologique du cytoplasme et du noyau ; ainsi qu'une classification connexionniste par le modèle du perceptron multicouche.

Les différentes étapes de ce système, se résument dans le dans le schéma suivant :

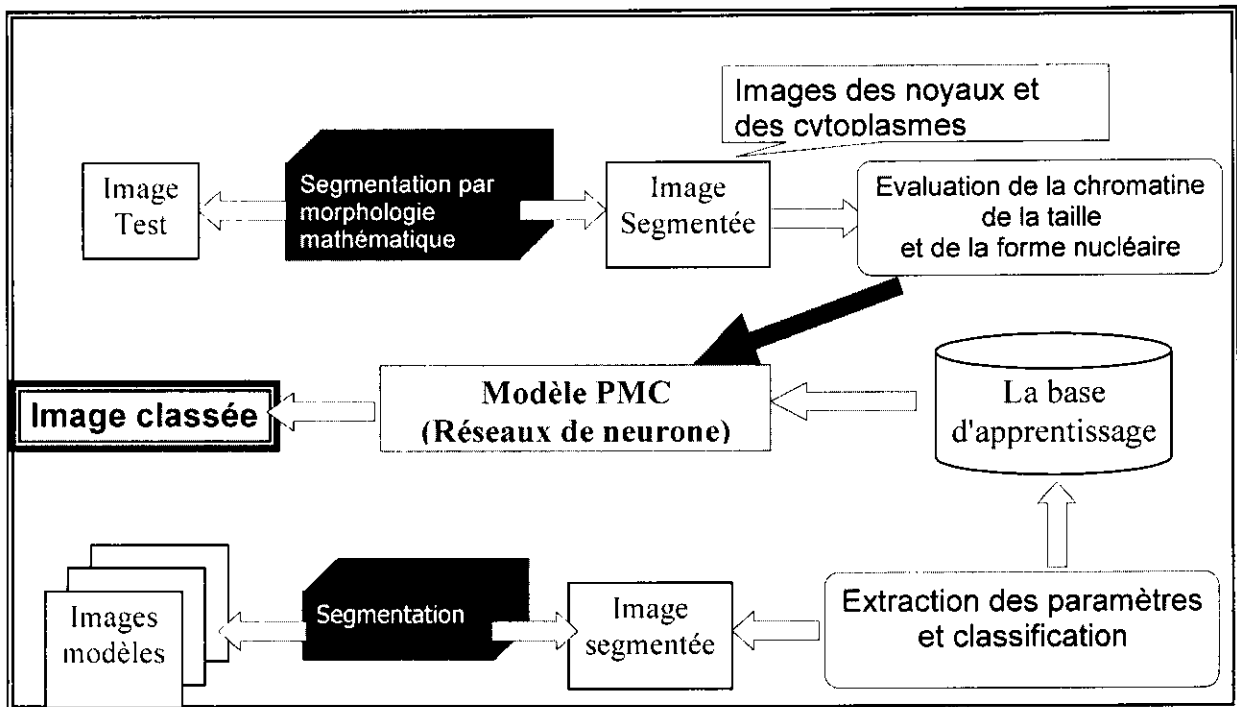


Figure 45 : Synoptique de l'application cytomorph

### II. 1 Interface du logiciel

La fenêtre principale de cytomorph est présentée par la figure suivante :

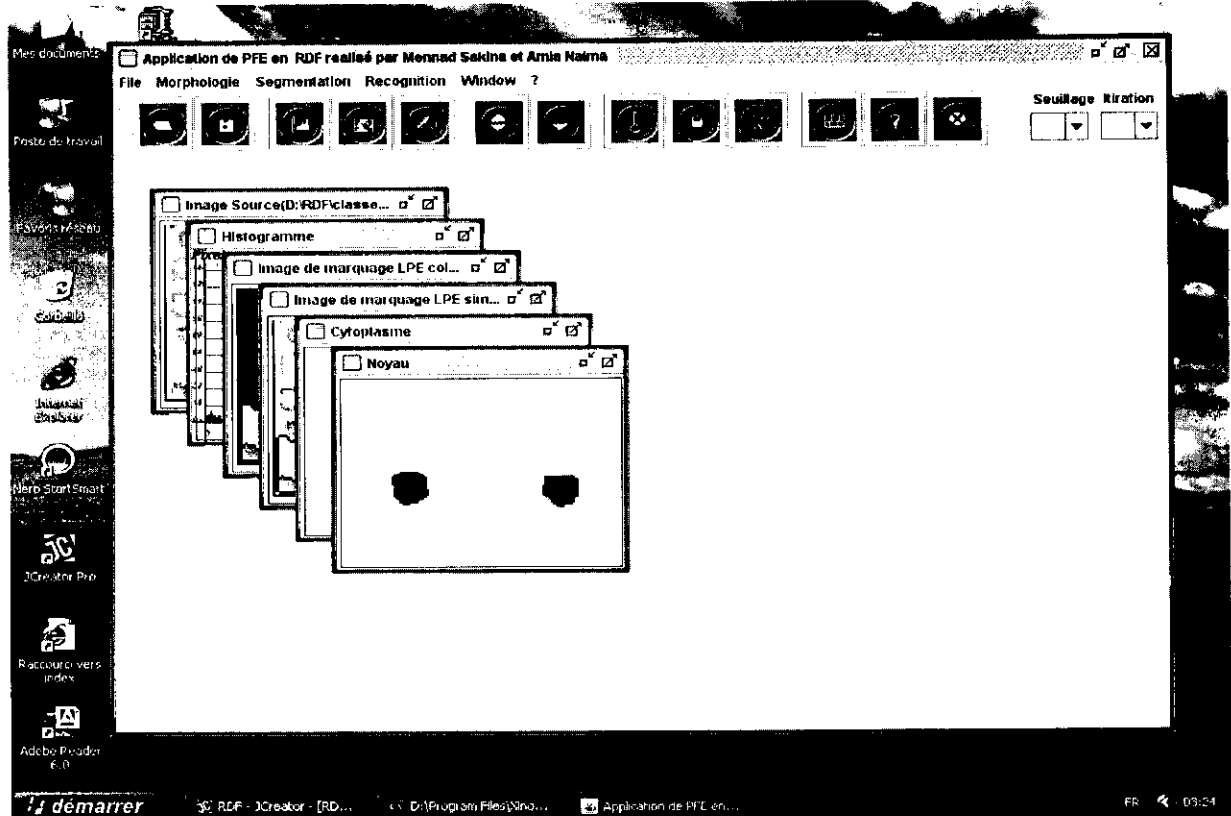


Figure 46 : Fenêtre principale de cytomorph

Nous apercevons la barre du menu principal, la barre d'outils et la barre d'état.

## II.2. Les menus

Dans la fenêtre principale on aperçoit la barre de menu qui englobe tous les opérations nécessaires pour notre application. Elles sont ils citées comme suit :

### II.2.1. Menu « File »

Nous trouvons dans ce menu, six commandes :

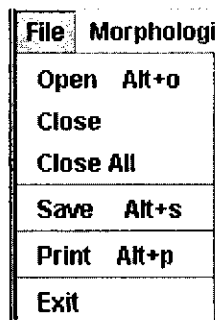


Figure 47 : Menu « fichier »

- **Open** : permet d'ouvrir le dossier contenant toutes les images PNG qu'on veut afficher,
- **Close** : permet de fermer la fenêtre sélectionnée,

- **Close all** : permet de fermer toutes les fenêtré qui existantes,
- **Save** : permet d'enregistrer l'image sélectionnée dans le dossier....
- **Print** : permet d'imprimer les images,
- **Exit** : sert à sortir complètement de la fenêtré principale.

**II.2.2. Menu « Morphologie »**

Ce menu contient six sous menus :

<b>Morphologie</b>
Delatation
Erode
Open
Close
OpenClose
Gradient

**Figure 48 : Menu Morphologie**

- **Dilatation** : permet d'exécuter l'opération morphologique de dilatation,
- **Erode** : permet d'exécuter l'opération morphologique d'érosion,
- **Open** : permet d'exécuter l'opération morphologique d'ouverture,
- **Close** : permet d'exécuter l'opération morphologique de fermeture,
- **Open Close** : permet d'exécuter l'opération morphologique d'ouverture et de fermeture,
- **Gradient** : permet de calculer le facteur gradient de l'image.

**II.2.3. Menu « Segmentation »**

Nous trouvons quatre sous menus :

<b>Segmentation</b>
Color Space ▶
Filtre ▶
LPE ▶
Detecte ▶

**Figure 49 : Menu Segmentation**

- **Color Space** : ce menu contient quatre sous menus; il sert à choisir l'espace couleur de l'image.

<b>Segmentation</b>	Reconition
<b>Color Space ▶</b>	RGB
Filtre ▶	HSL
LPE ▶	XYZ
Detecte ▶	L*U*V*

**Figure 50: Menu Color Space**

- ✓ **RGB** : permet d'afficher l'image dans l'espace couleur RVB,
  - ✓ **HSL** : permet d'afficher l'image dans l'espace couleur HSL,
  - ✓ **XYZ** : permet d'afficher l'image dans l'espace couleur XYZ,
  - ✓ **L\*U\*V\*** : permet d'afficher l'image dans l'espace couleur L\*U\*V\*.
- **Filtre** : ce menu contient deux sous menus ; il sert à choisir la méthode de filtrage pour l'amélioration de l'image.

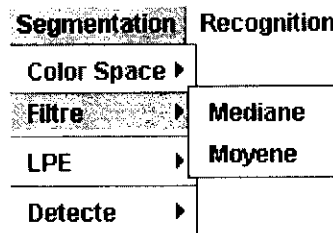


Figure 51: Menu Filtre

- ✓ **Médian** : permet de filtrer l'image par le filtre médian.
  - ✓ **Moyen** : permet de filtrer l'image par le filtre moyen.
- **LPE** : il sert à appliquer la méthode de la LPE amélioré a l'aide des trois sous menus suivant :

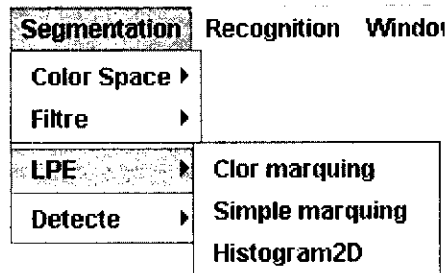


Figure 52: Menu LPE

- ✓ **Color marquing** : c'est un sous menu qui sert à segmenter les régions en couleurs.
  - ✓ **Simple marquing** : c'est un sous menu qui sert à segmenter les contours en noir.
  - ✓ **Histogram 2D** : c'est un sous menu qui sert extraire un histogramme à deux dimensions.
- **Détecte** : ce menu a deux sous menus qui sont :

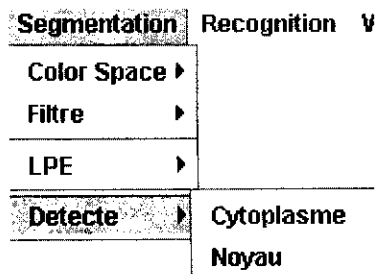


Figure 53: Menu Détecte



- **Cytoplasme** : permet d'afficher les régions cytoplasmiques,
- **Noyau** : permet d'afficher les région nucléaires.

#### II .2.4. Menu « Recognition »

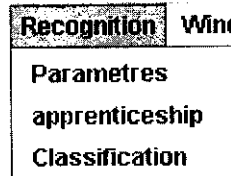


Figure 54: Menu Recognition

Le menu ci-dessus affiche trois sous menus qui sont :

- **Parametres** : il nous affiche les quatre critère de classification.
- **Apprenticeship** : il nous permet d'estimer les paramètres sous une base d'apprentissage.
- **Classification** : il nous permet d'afficher les résultats de la classification neuronale.

#### II.2.5. Menu Window

Le menu Window (Fenêtre) a deux sous menus :

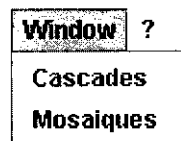


Figure 55: Menu Window

- **Cascades** : permet d'afficher les fenêtres en cascades,
- **Mosaïques** : permet d'afficher les fenêtres en mosaïques.

#### II.2.6. Menu « Help »

Nous observons dans le menu « Help » deux sous menus :

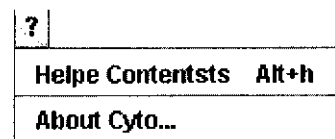


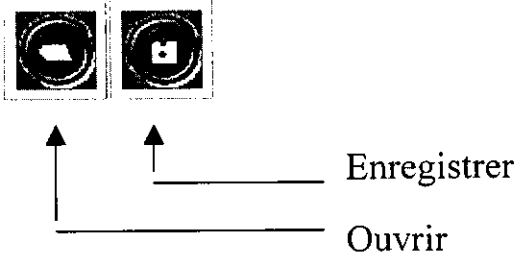
Figure 56: Menu Help

- **Help Contents** : permet de donner plus d'informations sur le contenu.
- **About Cyto** : permet d'afficher la fenêtre « A propos » du logiciel.

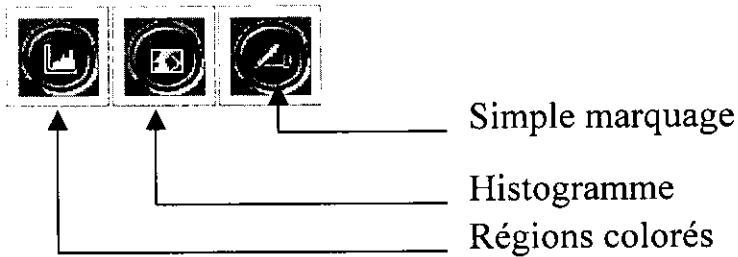
II.3. La barre d'outils

Nous avons organisé la barre d'outils selon le fonctionnement de la façon suivante :

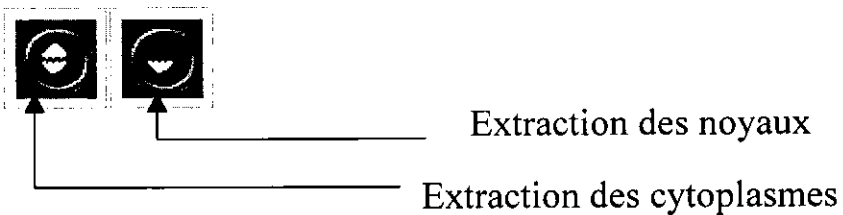
- La barre d'outils standard :



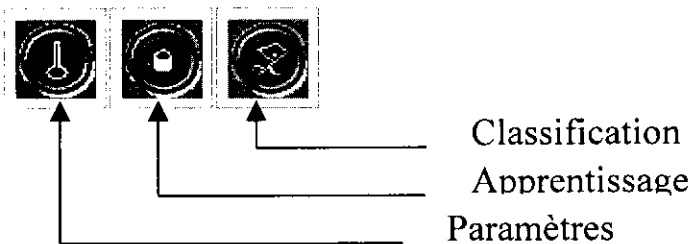
- La barre d'outils de LPE :



- La barre d'outils de l'extraction d'objet :



- La barre d'outils de classification :



- La barre d'outils auxiliaire :

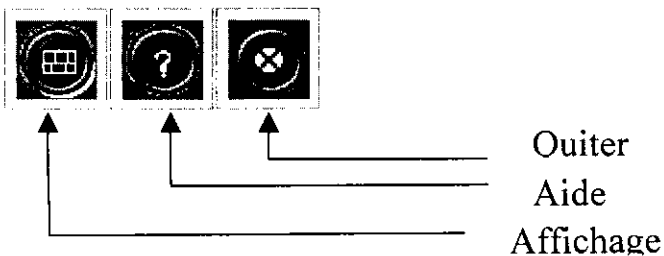


Figure 57 : Barre d'outils

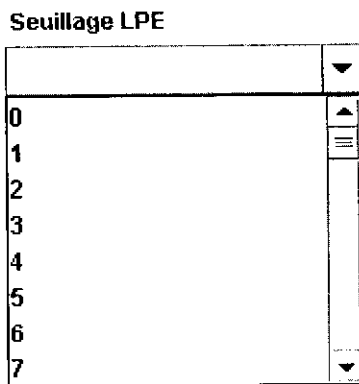


Figure 57: choix de seuil

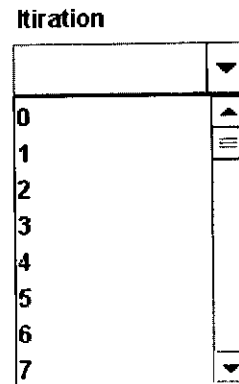


Figure 58: choix de membres d'itération pour les fonction morphologique

### III. Présentation des résultats et discussion

L'application cytomorph, permet d'analyser les images cytologiques, suivant une détection par la morphologie mathématique, une caractérisation par des paramètres morphologiques et une classification connexionniste pour l'aide à l'interprétation cytopathologiste.

Les étapes illustrant notre application sont les suivantes :

➤ **Première étape** : élimination du bruit.

Puisque l'image de gradient est très sensible au bruit, on effectue tout d'abord un filtrage ou un lissage afin d'éliminer le bruit sans trop altérer le signal. L'image de la figure 58.a, subit un filtrage gaussien; et un filtrage médian.

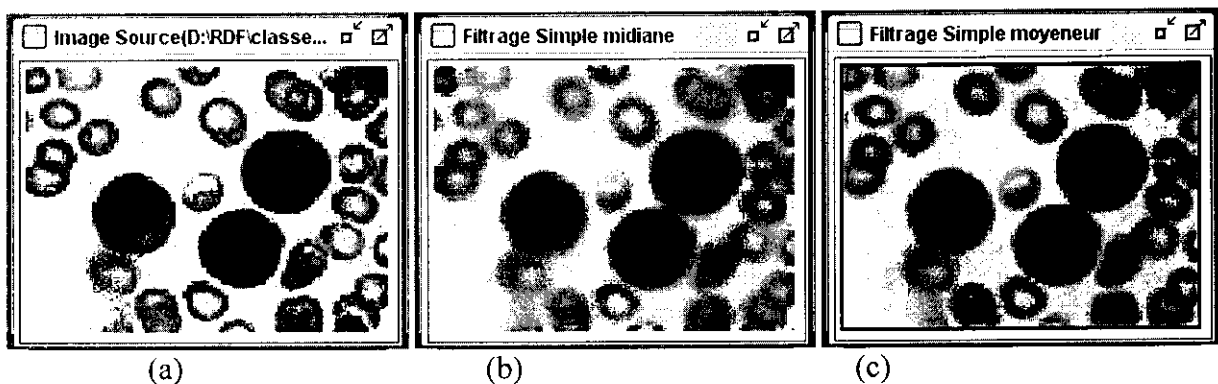


Figure 58: Résultats des différents filtrages : (a) image originale, (b) image filtrée par un filtre Médian, (c) image filtrée par un filtre moyen

➤ **Deuxième étape** : détection des cellules d'intérêt.

Pour détecter les cellules d'intérêt, l'élimination des globules rouges et d'autres corps est préférable. On ne garde que les cellules possédant un noyau bleu (cf. figure 59)

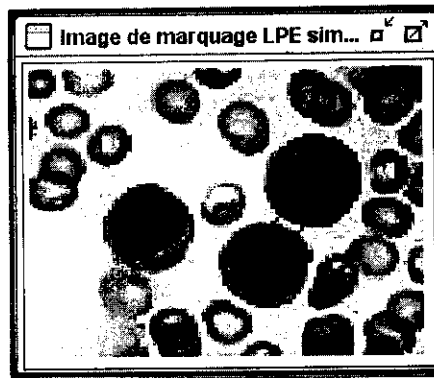


Figure 59: Détection des cellules d'intérêt

➤ **Troisième étape** : affichage de l'image couleur sous différents espaces couleurs.

Le choix d'un espace couleur, peut permettre d'améliorer les résultats d'une segmentation couleur. Parmi les espaces couleurs suivants : RVB, XYZ, LuVv, nous avons choisi le RVB, car il présente un meilleur contraste par rapport au autres (cf.figure 60).

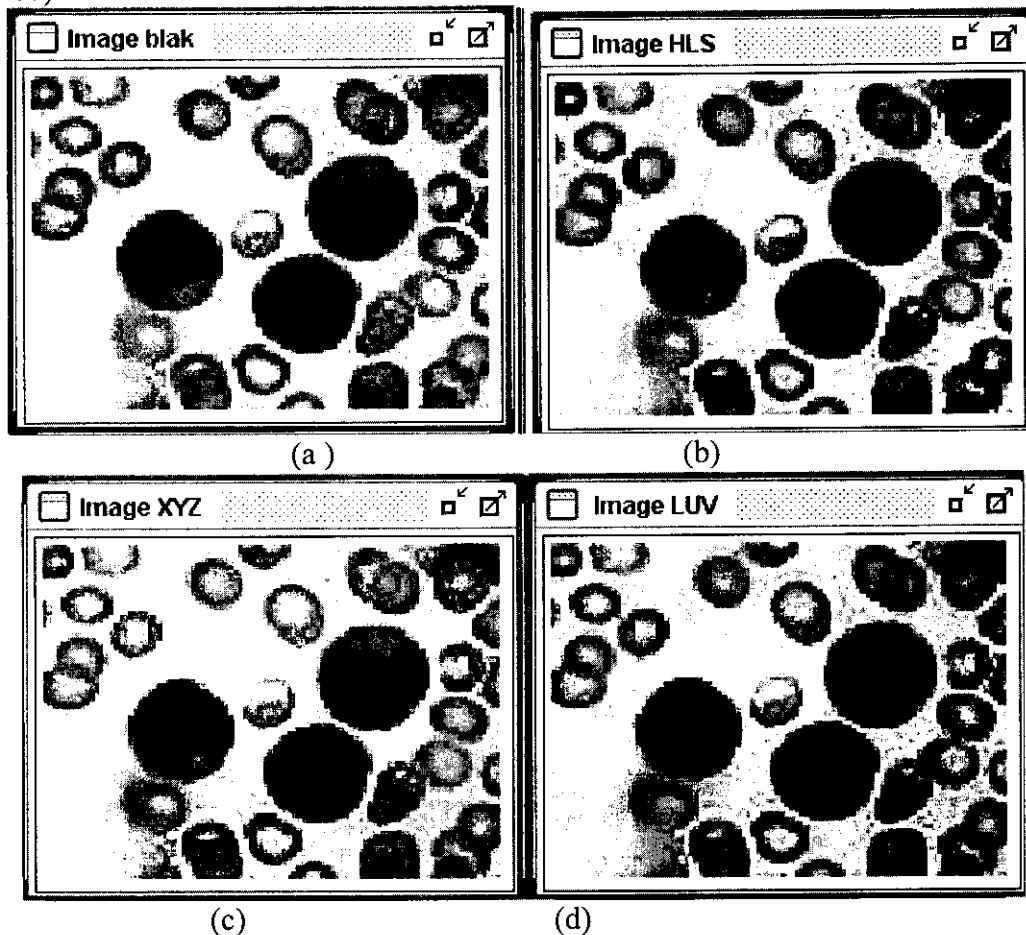
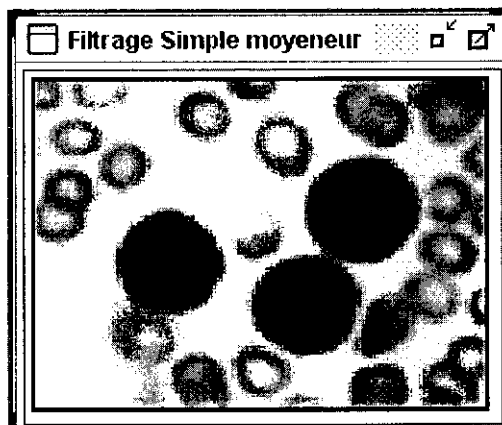


Figure 60: Représentation de l'image originale (12.a) dans différents systèmes couleurs ((a)image haute contraste, (b) image HSL, (c) image XYZ, image L\*U\*V\*)

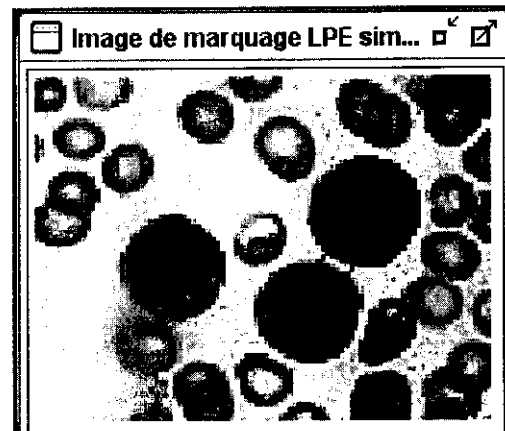
➤ **Quatrième étape** : la segmentation.

La segmentation morphologique par la LPE améliorée passe par les étapes suivantes :

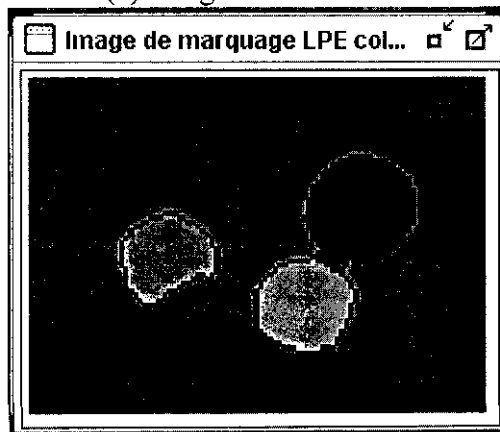
- ✓ Déterminer le gradient couleur : on constate que le gradient est élevé pour des points au voisinage desquels l'image, présente une forte discontinuité et faible pour les points au voisinage desquels l'image varie peu (cf. Figure 61.d).
- ✓ Déterminer l'histogramme 2D, où on constate des pics au niveau des intensités des couleurs les plus dominantes sur l'image par rapport au nombre de pixels (cf. Figure 61.e).
- ✓ Déterminer les différentes opérations de la morphologie mathématique (érosion, dilatation, ouverture, fermeture) afin de contraster ou dilater les objets (cf. Figure 61.f).
- ✓ Préciser les marqueurs cytoplasmiques, où le gradient permet d'extraire les régions cytoplasmiques (cf. Figure 61.f).
- ✓ Préciser les marqueurs nucléaires, où le gradient permet aussi d'extraire les régions nucléaires (cf. Figure 61.h).



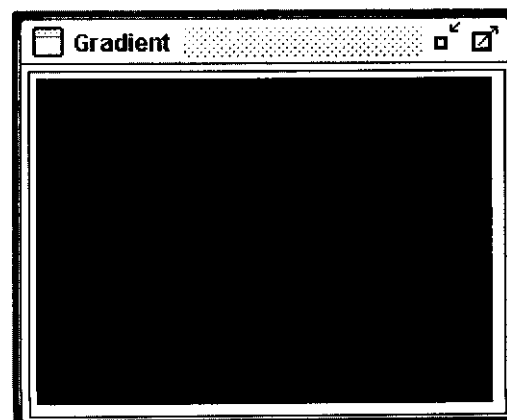
(a) Image filtrée



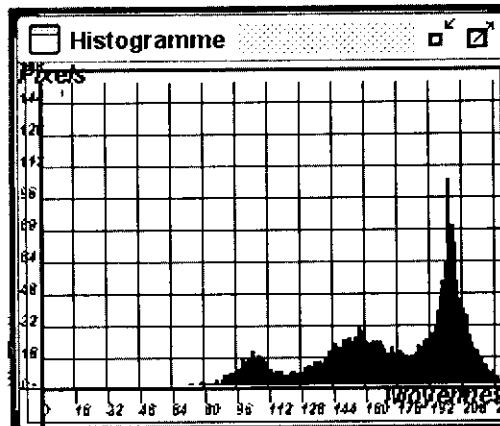
(b) Contours cytoplasmiques



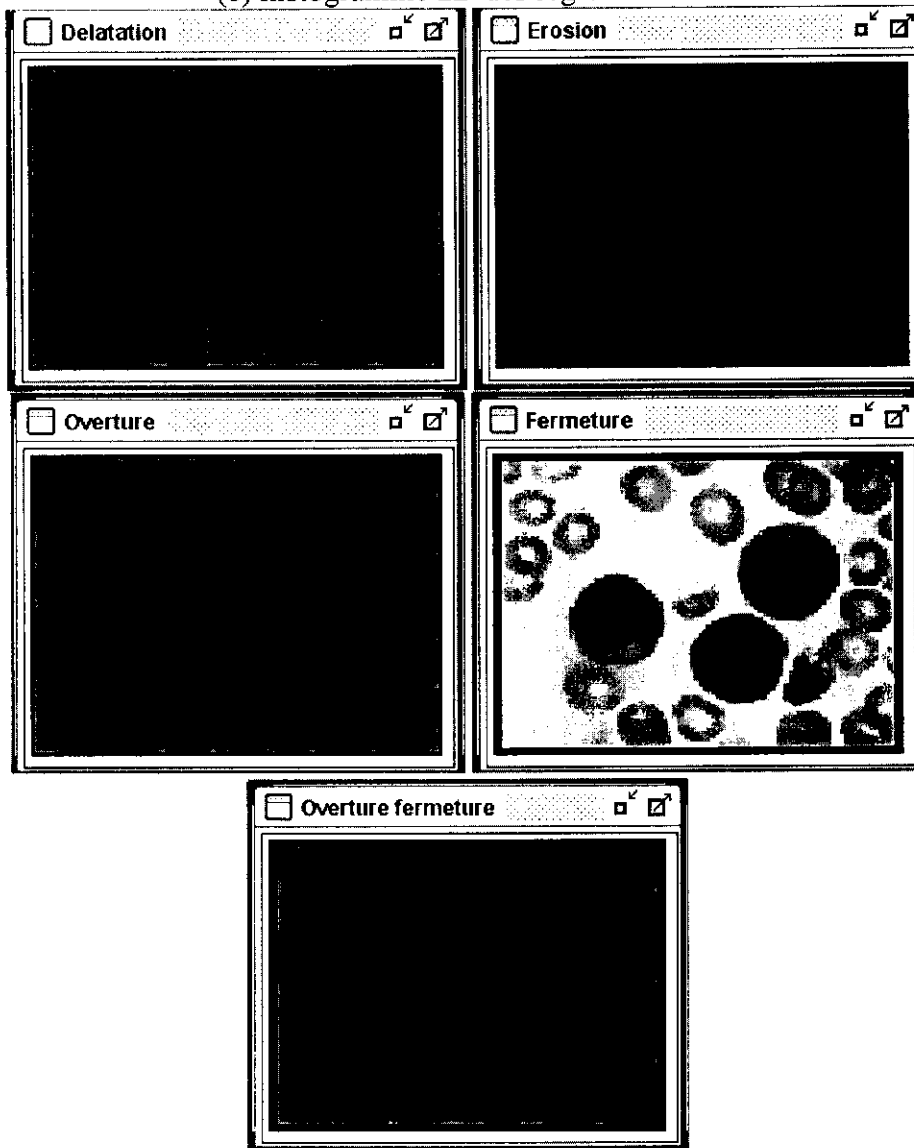
(c) Segmentation des régions en couleur



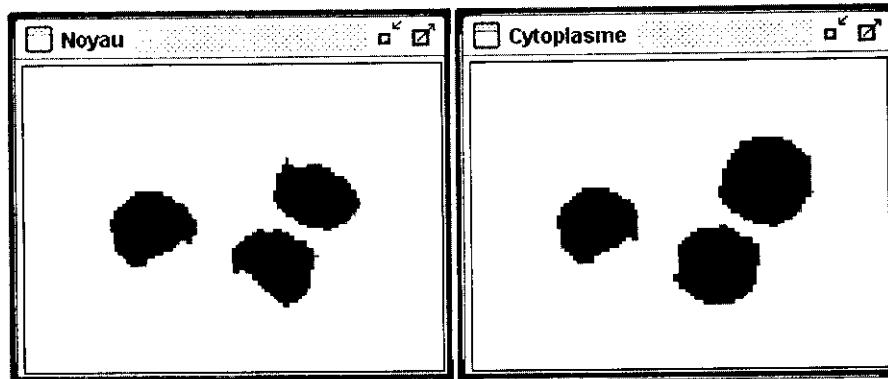
(d) Gradient couleur



(e) histogramme 2D des régions



(f) les opération morphologiques : dilatation, érosion, ouverture, fermeture et ouverture fermeture à la fois.



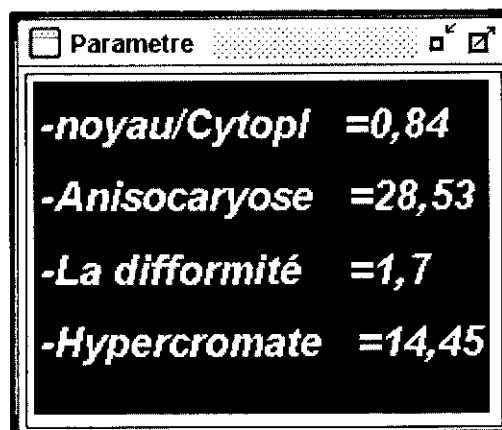
(g) extraction des cytoplasmes      (h) extraction des noyaux

**Figure 61:** Segmentation cytoplasmique et nucléaire avec précision du centre de gravité.

#### Cinquième étape : caractérisation des cellules cancéreuses.

Les quatre critères cytologiques permettant d'identifier la plupart des cellules cancéreuses sont les suivants :

- ✓ Le rapport nucléo-cytoplasmique, qui permet de quantifier le rapport des surfaces entre le noyau et le cytoplasme (cf. Figure 62),
- ✓ l'anisocaryose, qui permet de connaître l'irrégularité de la taille des noyaux d'une cellule à l'autre (cf. Figure 62),
- ✓ la difformité nucléaire, pour savoir le caractère non circulaire de la forme des cellules pathologiques (cf. Figure 62).
- ✓ l'hyperchromatisme nucléaire, qui permet de déterminer l'accroissement anormal de l'activité et la valeur de la texture apparente au sein des noyaux (cf. Figure 62).



**Figure 62 :** Résultats de calcul des paramètres pour l'image de la figure 58.a

#### ➤ Sixième et dernière étape : la classification.

Nous enregistrons les quatre paramètres précédents dans la base d'apprentissage (cf. Figure 63) puis on s'occupe de charger toutes les images (16 images) en mémoire afin qu'ils soient utilisés par le réseau de neurones lors de l'apprentissage. Notre stratégie de programmation a été la suivante : le neurone qui est l'entité la plus petite est modélisée en une simple structure

composée de quatre entrées (les quatre critères de malignité), trois sorties : la sortie 1 correspond à des cellules saines, la sortie 2 correspond à des cellules suspectes et la sortie 3 correspond à des cellules pathologiques) et ; de la liste des poids en rapport aux liaisons existant entre ce neurone et les neurones de la couche précédente.

L'apprentissage que nous utilisons, est un apprentissage automatique qui détermine l'architecture optimale du réseau à savoir son nombre de neurones de la couche cachée. Cela consiste à effectuer une suite d'apprentissages sur chacun des réseaux du système en utilisant, la méthode de **rétropropagation**. Ceci augmente à chaque itération, le nombre de neurones de la couche cachée. Le processus continue jusqu'à ce que la qualité obtenue, atteigne un certain seuil de qualité. Ainsi, chaque réseau détermine sa structure lors de l'apprentissage et sa qualité de classification est performante.

Le tableau ci-dessous, présente les valeurs des quatre paramètres morphologiques ainsi que la classification de 16 images cellulaires.

Numéro De l'image	rapport nucléocytoplasmique	anisocaryose	difformité nucléaire	hyperchromatisme nucléaire	Classe de l'image
1	0.89	0	9.92	18.14	pathologique
2	0.86	13.5	1.74	31.23	pathologique
3	0.98	6.5	1.85	12.08	pathologique
4	0.52	104.34	1.87	25.38	pathologique
5	0.80	146.51	1.29	41.2	pathologique
6	0.94	165	4.77	27.64	pathologique
7	0.60	20	6.68	30.14	pathologique
8	0.64	14	1.97	34.49	pathologique
9	0.60	22.2	2.69	21.9	pathologique
10	0.88	8	1.37	32.17	pathologique
11	0.94	31.62	1.81	71.41	pathologique
12	0.86	15	0.89	42.57	pathologique
13	0.62	25.53	1.65	29.31	pathologique
14	0.66	45.99	0.79	51.77	pathologique
15	0.50	7.5	1	34	suspecte
16	0.25	9.89	0.5	30.22	saine

**Tableau 2** : Classification de 16 images cellulaires à l'aide des quatre paramètres de malignité

Nous avons donc un tableau de 16 vecteurs sans oublier que :

- La classe 1 correspond à des cellules saines,
- la classe 2 correspond à des cellules suspectes.
- la classe 3 correspond à des cellules pathologiques.

Nous pouvons donc procéder à la classification de n'importe quelle image (cf. Figure 64), pour décider si elle appartient à une des trois classes choisies (classe 1 ou 2 ou 3).



Après avoir procédé à l'apprentissage du réseau Perceptron (cf. Figure 63), nous allons classifier par exemple l'image58.a selon les valeurs de critères de malignité (cf. Figure 64) :

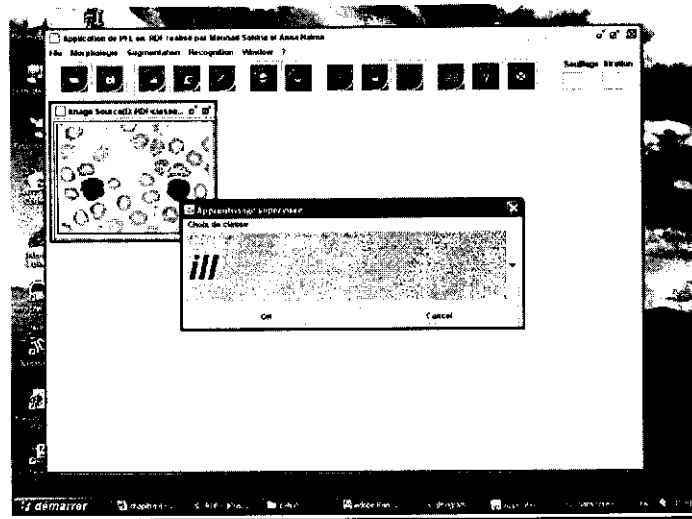


Figure 63: Apprentissage connexionniste par PMC

On constate dans cette figure, que le rapport nucléo-cytoplasmique égal à **0.84**, l'hyperchromatisme nucléaire égal à **14.45**, l'anisocaryose égale à **28.53** et la difformité nucléaire égale à **1.7**; paramètres supérieurs aux paramètres des cellules saines; l'image est donc classée parmi les cellule cancéreuses.

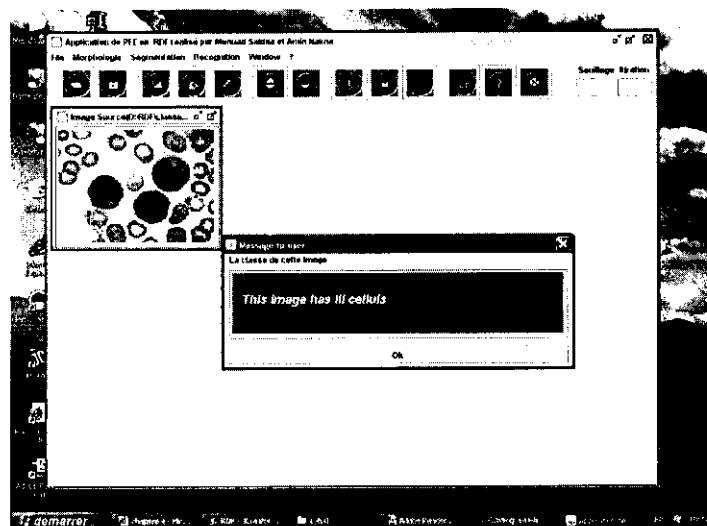


Figure 64 : Classification des images par le réseau PMC

Une classification connexionniste appelée "morphocyt" suivant le modèle de kohonen a été développée au cours d'un projet d'ingénieur de l'année 2006 [DJE 06], en utilisant uniquement deux paramètres pour la caractérisation ; à savoir l'hyperchromatisme nucléaire et le rapport nucléo-cytoplasmique risque de confondre les classes suspectes et les pathologiques. En effet, une comparaison de résultats de classification sur 3 images (cf.

Tableau 2) entre les deux applications : morphocyt et cytomorph, nous permet de remarquer une confusion entre la classe suspecte et la classe pathologique pour l'image de la figure 58.a

Nom d'application	Numéro d'image	Niveau de classe
Morphocyt	1	2
	2	3
	3	3
Cytomorph	1	3
	2	3
	3	3

**Tableau 3** : Comparaison des résultats entre deux types de classification

L'augmentation des paramètres permet de mieux analyser les images cytologiques.

#### IV. Conclusion

L'implémentation de notre méthode de segmentation morphologique, la LPE améliorée ainsi que la classification connexionniste par le réseau de neurone suivant le du perceptron multi couche, a engendré des résultats clairs, précis et encourageants.

L'application cytomorph est testée sur 13 images sélectionnées parmi une base de données de cellules de cas de leucémie. L'évaluation des résultats consiste à vérifier la bonne détermination des marqueurs cytoplasmiques et nucléaires, pour bien quantifier le rapport nucléo-cytoplasmique, l'anisocaryose, la difformité nucléaire et l'hyperchromatisme nucléaire; afin de classer les cellules suivant trois classes selon le modèle perceptron multi couches.

Ces résultats offre au pathologiste une aide à l'interprétation pour des images en cytologie.



# *CONCLUSION*

## Conclusion générale

Le cancer comme d'autres maladies chroniques, par exemple l'hypertension artérielle et ses conséquences, sont dues à une exposition à un environnement particulier avec cependant une différence notable dans le cas du cancer, la présence d'altérations génétiques causales.

L'objectif du traitement d'image, est d'aider le pathologiste dans son diagnostic. L'analyse du contenu des images selon plusieurs points de vue, nous amène à présenter notre projet de détection des cellules saines ou pathologiques suivant une classification connexionniste. Nous présentons la structure des différentes méthodes, ses implications, et les résultats obtenus.

Une des applications principales de la cytologie au sein d'un laboratoire d'anatomie pathologique, est d'aider le cytopathologiste à caractériser au mieux les cellules qui se multiplient de manière anarchique dans les cancers. Le projet pour lequel, nous avons eu un intérêt, est la détection de cellules cancéreuses à partir d'images cytologiques couleurs pour l'aide au diagnostic. Nous avons alors pu mettre en œuvre une méthode d'analyse d'images basée sur une segmentation par la ligne de partage des eaux améliorée, qui prend en compte l'information globale et locale suivant l'analyse de l'histogramme où nous nous sommes plus particulièrement intéressés au choix de l'espace couleur et à l'obtention de marqueurs. Notre méthode de segmentation permet ainsi, l'extraction de l'image des cytoplasmes et celle des noyaux dans le but d'une caractérisation des cellules pathologiques. Cette caractérisation, et faite autour de paramètres morphologiques prenant en compte la nature elle-même de la cellule.

La classification des cellules caractérisée, est faite suivant le modèle du perceptron. L'utilisation de réseau de neurone comme outil de classification est encore peu classique et, nous montrons en particulier que l'algorithme de perceptron multicouche s'avère un classifieur bien adapté pour les images couleurs. La méthode proposée, fournit de bons résultats.

Notre classification permet de répartir nos images cytologiques selon qu'elles sont composées de cellules saines, suspectes ou pathologiques.

Le processus développé peut être appliqué à divers cas de cellules cancéreuses. Le traitement d'images différentes, ne nécessitera qu'un simple ajustage des paramètres du système (taille et forme de l'élément structurant, nombre de neurones de la couche cachée dans le réseau connexionniste). L'extraction de plus de paramètres morphologiques des images cellulaires, participe à une caractérisation encore approfondie.

Quant, à la construction du classificateur, elle peut être totalement adaptée à l'utilisation que l'on veut en faire. Par exemple, un médecin intéressé uniquement par des coupes provenant d'un organe donné, utilisera des images d'apprentissage caractéristiques des cellules qu'il désire traiter. Le classificateur obtenu pourra donc être très pointu et spécialisé : il suffira de lui présenter des images spécifiques en grand nombre lors de la phase d'apprentissage. Si, de plus, les images sont très homogènes (conditions d'acquisition identiques, ...), les performances du classificateur seront encore accrues. On pourrait aussi, imaginer un praticien possédant plusieurs classificateurs, chacun étant adapté à un organe ou à un tissu particulier. Par contre, si l'on se choisit des images d'apprentissage très générales, provenant d'un ensemble d'organes différents, mais répondant bien sûr aux mêmes critères de malignité, on a alors la possibilité de construire un classificateur de type plus "généraliste".

## Conclusion générale

L'élaboration de notre système est donc réalisée, cependant certains points restent à approfondir. La base doit être complétée par davantage de cellules pour améliorer l'apprentissage et équilibrer les proportions des différentes classes de cellules existantes. Signalons enfin qu'en ce qui concerne l'apprentissage, chaque nouvelle image analysée lors de la phase de classification, pourra venir enrichir la séquence d'entraînement et permettra d'améliorer de façon continue la qualité de la classification.

Le traitement d'images n'a pas pour objet de remplacer les techniques traditionnelles utilisées pour le diagnostic des pathologies. La cytologie et l'histologie restent les techniques de référence pour la classification de ces entités. Toutefois, la cytogénétique et la biologie moléculaire permettent d'obtenir des informations complémentaires permettant de mieux définir ces entités et de fournir des éléments importants pour définir l'agressivité d'un processus néoplasique, de prédire la réponse à une thérapie particulière ou encore pour suivre l'évolution de la maladie au cours du traitement. La méthodologie étudiée est une porte ouverte aux systèmes futurs pour l'aide à la décision.

## Références bibliographiques

- [BEA 60] Beale L.S. "Examination of sputum from a case of cancer of the pharynx and adjacent parts", *Arch Med*, Vol22, p 44, 1860.
- [BEC93] M.O. Becks, Morphologie mathématique et quantification vectorielle pour la classification d'images médicales, mémoire de fin d'études, Université de Louvain, juin 1993.
- [BEU 90] Serge Beucher, « Segmentation d'image et morphologie mathématique », thèse de doctorat, école Nationale supérieure des Mines de Paris, 1990.
- [BON 06] Noël Bonnet, « Analyse multi variée des images multi- dimensionnelles », Université de Reims et UMRS Inserm 514,2006.
- [CIE 31] Commission Internationale de l'Ecléage, *Proceeding of the International congress on Illumination*, Cambridge university press, 1931.
- [COC 98] J.P Cocquerez et S.Philipp, « Analyse d'images : filtrage et segmentation » Edition Masson 1995.
- [COS 89] Machel Coster & J.L.Cherman, « Traitement d'image par morphologie mathématique application au domaine médical et industriel », Thèse de doctorat, université Blaise
- [DJE 06] Djeddi Omar, "Classification connexionniste d'image couleurs: Application aux images cytologiques, mémoire d'ingénieur, université de Blida, 2006.
- [DRE 02] Gérard DREYFUS « RECONNAISSANCE DE FORMES PAR RESEAU DE NEURONES » ECOLE POLYTECHNIQUE, France, juillet 2002. Pascal, 1987.
- [DUF04] Dufour H.M. "Meningite sarcomateuse diffuse avec envahissement de la moelle et des racines : cytologie positive et spéciale du liquide céphalo-rachidien," *Rev Neurol Soc Neurol*, Vol 12, p 204, 1904.
- [KUN 98] Kunt Murat, Gossta grand, Michel Kocher, « Traitement numérique de l'image », Lavoisier, Collection électricité traitement de l'information, 1998.
- [KOS 89] Koss L.G. "The Papanicolaou test for cervical cancer detection, a triumph and a tragedy", *J. Am. Med. Assoc.*, Vol 261, p737-743, 1989
- [LEZ 98] O. Lezoray, A. Elmoataz, H. Cardot, G. Gougeon, M. Lecluse, H. Elie, M Revenu, « Segmentation d'Images Couleur : application à la cytologie des séreuses en vue d'un tri informatique cellulaire », 1998.
- [LEZ 95] Olivier Lezorray « La réalisation d'une étude sur le tri informatique en cytologie », Rapport de stage DEA, université de Caen France, 1995.
- [LEZ 00] Olivier LEZORAY, Abderrahim ELMOATAZ, Hubert CARDOT, Marinette EVENU, « Segmentation d'images couleur : applications en microscopie cellulaire », 2000.

[LEG 02] Christophe LÉGER, « Traitement des Images », ESPEO, 2002-2003.

[LEZ 02] Olivier Lezorray « segmentation d'image couleurs par morphologie mathématique et classification des données par réseaux de neurones : application a la cellules en cytologie des séreuses », Thèse de doctorat, université de Caen France, 2002.

[LEZ 04] Olivier Lezorray «An unsupervised color image segmentation based in morphological 2D clustering and fusion »,2004.

[MAR 09] Marissi G. "Ueber die diagnose des magencarcinomas auf Grund der Cytologie des Spuelwasser", *Arch Verdauungskrankh*, vol 15, p 251, 1909.

[Mey 92] F. Meyer, « Color Image Segmentation », Proceedings of the ICIP,pp. 303-306, 1992.

[MUEL 38] Mueller J. "Ueber den feineren Bau und die Formen der krankerhaften geschwuelste", *G Reimer*, Berlin 1838.

[PAP 47] Papanicolaou G.N. Traut HF. "Diagnosis of Uterine Cancer by the Vaginal Smear" , *Commonwealth Fund*, NewYork, 1947.

[PAP 54] Papanicolaou G.N. "Atlas of Exfoliative Cytology", *Harvard University Press*, Cambridge, Massachusetts,1954.

[PER 90] B. Péroche, J. Argence, D. Ghazanfarpour, D. Michelucci « Traité des Nouvelles Technologies » – HERMES – 1990

[RAC 02] Victor Racine, Mémoire de fin d'études, Institut Curie, Mars – Septembre 2002

[RAM 05] Ramoul Amel et Boukrif Souad,"Segmentation d'image couleur: Application aux images cellulaires", mémoire d'ingénieur, université de blida, 2005.

[REV 67] "Revision of laboratory requisitions for the medicara, medicaid and clinical Laboratories improvement act of 1967 ( CLIA'67) programs", *Federal Register*, Vol 5, p 9600,1967

[SAN 64] Sanders W.R. "Cancer of the bladder : fragments forming urethral plugs discharged in the urine", *Med J*, Edinburgh, Vol 1, p 273, 1864.

[SER 82] J.Serra « Image analysis and mathematical morphology », academic Press,1982.

[SER 00] J.Serra « Morphologie mathématique », Ecole des Mines de Paris, 2000.

[YZI 05] YZIQUEL Jérôme et ERMONT Nicolas, « Apprentissage par Réseaux de neurones », ENSICAEN de France, année 2004-2005.

- [1] <http://www.info.unicaen.fr/~lezoray>
- [2] <http://www.greyc.ismra.fr>
- [3] <http://www.meurie.info>
- [4] <http://ligim.univ-lyon1.fr/textes/equipes/IMAGE/CORESA03/articles/13.pdf>
- [5] <http://www.neurones.espci.fr/Français.doc/livre.htm>
- [6] <http://www.polymorphe.org/telecharger-278-cours-Reseaux-de-Neurones.pdf>
- [7] <http://www.Developpez.com>.

