

Remerciement

Ce travail a été dirigé par Monsieur BENAZZOUZ Mourtada, Maître de conférences classe B et professeur d'informatique à l'université de Tlemcen. Je lui adresse à cette occasion mes sincères remerciements pour m'avoir encadré. Je tiens aussi très honnêtement, à lui exprimer mes reconnaissances pour avoir énormément appris sur les différentes méthodes de segmentation durant le semestre préparatoire. Ce sont au fait sa pédagogie, son savoir-faire, sa passion pour la recherche, sa vocation pour l'encadrement, sa sagesse et son dévouement pour la formation qui m'ont encouragés à travailler sous sa direction.

Je remercie également Dr.Benmansour Nadia et Mlle Benghebrit Houria pour avoir accepté de diriger mon travail, ainsi pour leur aide et leurs précieuses orientations à tout bout de chemin de mon stage. Sans oublié toutes les personnes du laboratoire centrale qui sans eux ce travail n'aurait pu être mené a bien

Mes remerciements s'adressent aussi aux membres de jury qui nous ont fait le grand honneur d'évaluer ce travail.

Dédicace :

Remerciements et louanges à Dieu, de m'avoir donné la foi et la force pour accomplir ce travail. Prière et salut sur notre prophète " Mohammed " et sur sa famille et ces
compagnons

Avec tout mon amour éternel et avec l'intensité de mes émotions. Je dédie ce mémoire :
à mes chers parents qui ont le droit de recevoir mes chaleureux remerciements pour le
courage et le sacrifice qu'ils ont consenti pendant la durée de mes études en leurs
souhaitant une longue vie pleine de joie et de santé.

Pour la même occasion je tiens à remercier ma seule et unique sœur Amina que je ne
cesserai jamais d'aimer, je te dédie ce travail avec tous mes vœux de bonheur, de santé
et de réussite.

Aux personnes les plus chères à mes yeux qui m'ont soutenu tout au long de mon
travail, pour qui ma réussite est très importante, mes cousins Imed, Adel, Zakaria et à
mes belles cousines Hidayet et Bouchra, ainsi que tous mes oncles et tantes, que Dieu
les protège et les garde pour moi.

A celui qui m'a soutenu et encourager tous long de la réalisation de ce projet : Abbou
Zoubir je le remercie pour m'avoir aidé pendant cette année.

A mes très chères grands mères Madjida et Fatima, que Dieu les protèges et nous les
garde.

À tous ceux qui me sont chers, si mes vœux pouvaient avoir quelques pouvoirs, j'en
serai profondément heureuse car je veux pour vous et vos familles toutes les réussites et
satisfactions de ce monde.

Table des matières

Table des figures	vii
Liste des tableaux	viii
Notations et Abréviations	ix
Introduction	x
1 Cytologie descriptive	1
1.1 Introduction	1
1.2 Hémogramme	2
1.2.1 Définition	2
1.2.2 Numération formule sanguine	2
1.2.3 Frottis sanguin périphériques	3
1.3 Myélogramme	5
1.3.1 Définition	5
1.3.2 Indications	5
1.3.3 Lecture du myélogramme au microscope	6
1.3.4 Valeurs normales	7
1.4 Coloration	7
1.5 Les éléments figurés du sang périphérique	8
1.5.1 Les globules rouges (Erythrocytes ou Hématies)	8
1.5.2 Les globules blancs (ou leucocytes)	9

1.5.3	Les plaquettes ou Thrombocytes	11
1.6	Conclusion	11
2	Traitement d'images et espace couleur	12
2.1	Introduction	12
2.2	Le pixel	12
2.3	Généralités sur l'image	13
2.3.1	Définition d'une image	13
2.3.2	Typologie d'image	14
2.3.3	Texture de l'image	14
2.3.4	Technologie de l'image numérique	15
2.3.5	Image noir et blanc	16
2.3.6	Image au niveau de gris	17
2.3.7	Résolution d'une image	17
2.3.8	Taille d'une image	18
2.4	Définition des grandeurs chromatiques	18
2.5	Exploitation de la couleur dans une image couleur	19
2.6	Espaces colorimétriques	19
2.6.1	Les espaces de primaires	21
2.6.2	Les espaces de luminance-chrominance	23
2.6.3	Les espaces perceptuels	24
2.6.4	Les espaces d'axes indépendants	26
2.7	Conclusion	27
3	Etat de l'art sur la segmentation d'images	28
3.1	Introduction	28
3.2	Problématique classification-segmentation	29
3.2.1	Classification	29
3.2.2	Segmentation	29
3.3	Différentes approches de la segmentation d'images	30

3.3.1	Approche contour	30
3.3.2	Approche région	30
3.4	Segmentation basée sur les pixels	36
3.5	Seuillage global	37
3.5.1	Les méthodes non paramétriques	37
3.5.2	les méthodes paramétriques	38
3.6	Conclusion	39
4	Présentation de l'application	41
4.1	Préliminaire	41
4.2	Etat de l'art	41
4.3	Base d'images	42
4.4	L'algorithme d'Otsu	42
4.4.1	Principe de fonctionnement	42
4.4.2	Implémentation	43
4.5	L'algorithme de Croissance de région	45
4.5.1	Principe de fonctionnement	46
4.5.2	Implémentation	48
4.5.3	Les paramètres de La fonction	49
4.5.4	Avantages et inconvénients de la méthode	51
4.5.5	Solutions proposées	53
4.6	Interface de l'application	54
4.6.1	Discussion du premier résultat	58
4.6.2	La phase d'hybridation	60
4.6.3	Evaluation de la segmentation	63
4.7	Conclusion	67

Table des figures

1.1	Les lignées sanguines	1
1.2	Phases de préparation de frottis sanguin	4
1.3	Frottis médullaire de richesse normale en grossissement 10	6
1.4	Hématopoïèse	8
1.5	Les types de globules blancs.	9
2.1	Image de texture couleur de la base Outek	15
2.2	Image numérique	16
2.3	Le pixel encadré en blanc de composantes couleur sont C1, C2 et C3 donne naissance au point C	20
2.4	Les couleurs des pixels des images donnent naissance aux nuages de points.	20
2.5	Le cube des couleurs.	22
2.6	Le diagramme de chromaticité (r_c , g_c) de la CIE	22
2.7	Représentation des fonctions colorimétriques R,V,B	23
2.8	Le diagramme de chromaticité (x , y) de la CIE.	24
2.9	Le système HSV de cône hexagonal simple (a) et double (b).	25
2.10	Le cube des couleurs dans le modèle triangulaire.	26
3.1	Découpage par quadtree d'une image 8x8.	33
3.2	Quadtree.	34
3.3	Problème de découpage arbitraire de régions dans la phase de division.	35
3.4	Principe de segmentation watershed.	36

4.1	Fonction d'Otsu.	44
4.2	Deux schémas de regroupement pour des pixels a, b et c.	52
4.3	Exemple de sous segmentation avec un seul point d'amorce.	53
4.4	Interface de l'application.	54
4.5	Extraction originale des noyaux.	55
4.6	Elimination du bruit.	56
4.7	Phase de remplissage.	56
4.8	Noyaux convexes et dilatés.	57
4.9	Le résultat de la croissance de région (1er cas).	59
4.10	Le résultat de la croissance de région (2ème cas).	59
4.11	Résultat de la fonction hybridée.	62
4.12	Comparaison et similarité entre le résultat et l'étiquète.	64
4.13	Classification et comparaison.	66
4.14	Les résultats de Classperf.	66

Liste des tableaux

1 valeurs normales de l'hémogramme chez l'adulte	5
2 Valeurs de référence des cellules de myélogramme chez l'adulte. . . .	7

Notations et Abréviations

PNN : polynucléaires neutrophiles .

PNE : polynucléaires éosinophiles.

HSV : Hue, Saturation, Value.

RVB Rouge, Vert, Bleu.

HSI : Hue, Saturation, Intensity.

RG : Region Growing

Introduction générale

Faisant appel à de nombreuses technologies, l'imagerie médicale est de plus en plus utilisée pour le diagnostic de nombreuses maladies en complément d'un examen clinique et d'autres investigations. En matière d'investigation par les examens biologiques, l'imagerie médicale demeure cet élément qui s'impose dans le monde de la médecine moderne. Elle permet aux spécialistes de poser leur diagnostic de manière rigoureuse, les aidant à la mise au point de nouveaux traitements et au développement de leurs recherches.

La pratique de l'étude de la morphologie cellulaire sanguine a beaucoup évolué ces derniers temps. Cette évolution est due, d'une part à l'apparition d'automates modernes d'hématologie utilisés dans l'identification de cellules normales circulantes par une technique de cytométrie en flux. D'autre part la microscopie optique demeure toujours importante, voire indispensable à la reconnaissance de cellules pathologiques sur frottis sanguin.

Une bonne connaissance de la morphologie d'éléments normaux du sang tels, globules rouges, plaquettes et leucocytes, reste alors un pas fondamental conduisant à l'étude de la pathologie. La médecine, bénéficie donc pleinement de tous les progrès des sciences de l'informatique, notamment dans l'analyse intelligente qui porte aide à la décision.

Cette analyse joue donc un rôle primordial qui facilite l'interprétation et la quantification des images de microscopies. La qualité et la robustesse des résultats d'analyse d'images dépend essentiellement de l'étape dite de « segmentation » qui consiste à extraire des zones ou objets d'intérêt.

La segmentation d'image est une étape critique de l'analyse d'images puisqu'elle

recouvre une multitude d'algorithmes. Ce qui engendre de véritables difficultés dans l'identification d'algorithmes utiles et efficaces avec l'élimination de ceux dont l'utilisation n'est pas adaptée ou potentiellement sources d'artefacts.

Notre travail consiste à mettre au point une nouvelle méthode de segmentation d'images cytologiques couleurs dont la configuration spatiale et colorimétrique "extrêmement variable", rend difficile l'extraction des composants.

Le principale objectif de ce travail se situe dans ce cadre. Nous nous proposons alors d'intégrer un indice de circularité dans la méthode de croissance de région. Cette méthode, appelée aussi *Region Growing* est l'une des méthodes les plus utilisées dans ce domaine médical.

Notre partie théorique porte sur la cytologie descriptive en tant que terrain d'études ; les caractéristiques et les différents espaces couleur de l'image que nous traitons. Nous présentons aussi par la suite les différentes méthodes de segmentation et de classification tout en abordant le choix des méthode de segmentation et du seuillage.

La partie expérimentale est consacrée à la présentation du principe de fonctionnement et à l'implémentation des méthodes utilisées. Elle est consacrée aussi à la mise au point de notre application.

Dans la partie résultats et discussions de ce mémoire, des images traitées ont été présentées, suivis d'une évaluation de la segmentation.

Enfin notre travail est terminé par une conclusion générale ouvrant le chemin aux perspectives et proposant un ou plusieurs points d'études afin d'assurer la continuité dans la recherche pour le développement des techniques de segmentations notamment dans le domaine de la cytologie.

Chapitre 1

Cytologie descriptive

1.1 Introduction

Dans ce chapitre nous présentons des aspects principaux utiles pour la compréhension de la cytologie descriptive en définissant l'hémogramme ainsi que le myélogramme. Avec les types des différentes cellules sanguines et leurs caractéristiques.

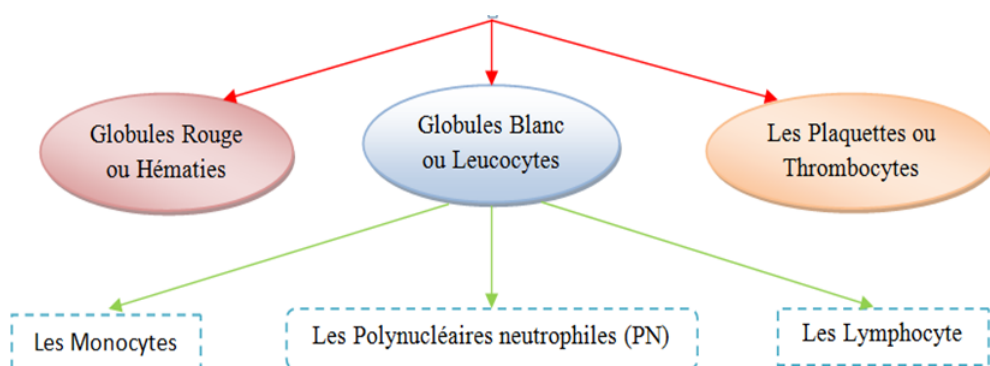


FIGURE 1.1 – Les lignées sanguines

1.2 Hémogramme

1.2.1 Définition

L'hémogramme est un examen courant en pratique, simple, peu coûteux et automatisé. Il repose sur l'étude quantitative et qualitative des éléments figurés du sang.

L'hémogramme consiste à :

- Faire la numération des cellules sanguines circulantes
- Calculer les constantes hématimétriques : VGM, CCMH, TCMH
- Établir la formule leucocytaire
- Étudier la morphologie des différentes lignées cellulaires (par le frottis sanguin périphérique : FSP)

1.2.2 Numération formule sanguine

A. Définition

Elle représente l'examen quantitatif qui consiste en un comptage de tous les éléments figurés du sang.

- Nombre de globules rouges
- Taux d'hémoglobine
- Hématocrite
- Constantes hématimétriques

B. Indications de l'hémogramme Il est indiqué lorsqu'il existe des pathologies hématologiques ou infectieuses comme :

- La PATHOLOGIE MYELOÏDE
- La PATHOLOGIE LYMPHOÏDE

1.2.3 Frottis sanguin périphériques

A. Définition

Le frottis sanguin est une étude morphologique des éléments figurés du sang étalés en couche monocellulaire sur une lame de verre après coloration.

Le frottis sanguin périphérique confirme la formule automatique, permet d'apprécier la morphologie érythrocytaire (taille, forme, coloration, inclusion ...), apprécier la population leucocytaire (granulation, segmentation des PNN...) et détecter une éventuelle myélémie, agrégat plaquettaire ou cellule anormale.[1]

B. Indications

Malgré la généralisation des automates capables de réaliser des formules leucocytaires, l'interprétation au microscope des éléments sanguins sur un frottis reste indispensable dans plusieurs circonstances :

- présence d'anomalies quantitatives de l'hémogramme touchant une ou plusieurs lignées, même en l'absence d'alarmes spécifiques ;
- difficultés d'identification des populations leucocytaires par l'automate, signalées par des alarmes,
- présence d'une anémie, en particulier pour examiner la morphologie des hématies, même si la formule leucocytaire proprement dite réalisée par l'automate est satisfaisante.[2]

C. Confection de frottis sanguin

L'étalement est réalisé à partir de sang frais (prélèvement de moins de 3 heures) prélevé sur éthylène-diamine-tétra-acétique (EDTA) ou, plus rarement, directement par prélèvement capillaire[3].

Dans notre stage au service d'hémobiologie dans le laboratoire de cytologie qui a duré une quinzaine de jours, nous avons appliqué la technique manuelle d'étalement.

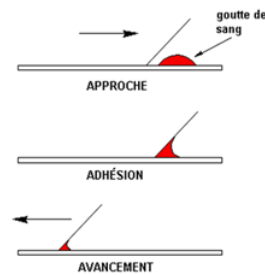


FIGURE 1.2 – Phases de préparation de frottis sanguin

[4]

- Nous avons commencé par déposer une gouttelette de sang près de l'extrémité d'une lame en verre, dégraissée et à pans coupés ;
- Ensuite nous avons appliqués une autre lame inclinée à 45 degrés en avant de la goutte de sang de façon à ce que le sang s'étale sous la lame par capillarité ;
- Après on a fait glisser la lame inclinée à 45 degrés pour étaler uniformément la goutte ;
- Enfin nous l'avons séché par agitation pour fixer les cellules(voir figure1.2).

D. Qualité du frottis

La bonne qualité du frottis est un préalable indispensable à l'analyse du frottis de sang après coloration. En effet, des frottis mal faits risquent d'être à la source d'erreurs dues à :

- une mauvaise appréciation du pourcentage des populations leucocytaires liée à une répartition hétérogène ;
- une difficulté d'identification des cellules anormales ;

E. Examen du frottis de sang

L'examen du FSP permet d'établir une formule leucocytaire et d'examiner la morphologie des hématies, des globules blancs et des plaquettes.

Valeurs normales d'un hémogramme

Lignée	Paramètre	Valeurs normales
Lignée érythrocytaire	Globules rouges (T/L)	Homme : 4,2-5,7 Femme : 4-5,3
	Hémoglobine (g/dl)	Homme : 13-18 Femme : 12-16
	Hématocrite (%)	Homme : 40-52 Femme : 37-46
	VGM (fL)	80-100
	TCMH (pg)	27-31
	CCMH (%)	32-36
Lignée leucocytaire	Leucocytes (G/L)	4-10
	Polynucléaires neutrophiles (G/L)	2-7,5
	Polynucléaires éosinophiles (G/L)	0,1-0,5
	Polynucléaires basophiles (G/L)	0-0,15
	Lymphocytes (G/L)	1,5-4
	Monocytes (G/L)	0,2-1
Lignée plaquettaire	Plaquettes (G/L)	150-400

Table 1 - valeurs normales de l'hémogramme chez l'adulte.[3]

1.3 Myélogramme

1.3.1 Définition

Le myélogramme représente l'un des principaux outils de diagnostic de la plupart des hémopathies.

Cette analyse consiste à prélever par aspiration quelques gouttes de suc médullaire riche en cellules hématopoïétiques et de les étaler par frottis sur des lames de verre. Après coloration, les cellules sont analysées au microscope par un cytologiste expérimenté, qui fournit un décompte de la répartition des cellules ainsi qu'une appréciation qualitative de leurs éventuelles anomalies cytologiques. Cet examen permet d'élaborer et/ou de contribuer au diagnostic de nombreuses hémopathies (leucémies aiguës, syndromes myélodysplasiques, syndromes myéloprolifératifs, myélome, lymphomes, etc.).

1.3.2 Indications

Le myélogramme permet de détecter trois principales catégories d'anomalies :

- présence de cellules anormales (leucémies, lymphomes par exemple) ;
- anomalies quantitatives de répartition des lignées hématopoïétiques. Devant une cytopénie, le myélogramme permet de trancher entre une étiologie centrale ou périphérique,

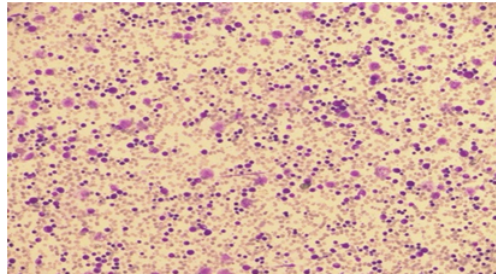


FIGURE 1.3 – Frottis médullaire de richesse normale en grossissement 10

[5]

selon que la lignée concernée est quantitativement diminuée ou non dans la moelle ;

- anomalies qualitatives concernant la morphologie cellulaire (syndromes myélodysplasiques, carences vitaminiques ou autres causes de mégalo blastose, par exemple).

En dehors d'anomalies de l'hémogramme, le myélogramme fait aussi partie du bilan d'autres anomalies telles qu'un pic monoclonal, une splénomégalie ou le bilan d'extension médullaire d'une tumeur solide.

1.3.3 Lecture du myélogramme au microscope

La lecture du myélogramme doit être effectuée par un cytologiste expérimenté et disposant d'un microscope de bonne qualité (100W).

On réalise une formule complexe en comptant, tout d'abord le pourcentage respectif de chacune des trois lignées de l'hématopoïèse : érythroblastique, myélocytaire et lymphoïde puis le pourcentage de chacun des différents stades dans chaque lignée.

1.3.4 Valeurs normales

Valeurs de référence (%)	
Richesse	Normale
Lignée mégacaryocytaire	Abondance normale
Blastes indifférenciés	0-2
Lignée érythroblastique :	
- Proérythroblastes	- 0,5-2
- érythroblastes basophiles	- 1-3
- érythroblastes polychromatophiles	- 5-15
- érythroblastes acidophiles	- 5-10
Lignée granuleuse :	
- myéloblastes	- 0-2
- promyélocytes	- 1-4
- myélocytes neutrophiles	- 10-15
- myélocytes éosinophiles	- 0-1
- métamyélocytes neutrophiles	- 10-20
- métamyélocytes éosinophiles	- 0-1
- polynucléaires neutrophiles	- 15-25
- polynucléaires éosinophiles	- 0-1
- polynucléaires basophiles	- 0-1
Autres cellules :	
- Lymphocytes	- 5-20
- Monocytes	- 0,5-2
- Plasmocytes	- 0-2

Table 2 - Valeurs de référence des cellules de myélogramme chez l'adulte.[5]

1.4 Coloration

La coloration repose sur l'action combinée de deux colorants neutres :

1. **le May Grunwald** : contenant un colorant acide : l'éosine et un colorant basique : le bleu de méthylène,
 2. **le Giemsa** : contient aussi de l'éosine et un colorant basique : l'azur de méthylène,
- Ces deux colorants sont en solution dans l'alcool méthylique (fixateur).

En période du stage préalablement cité, on a appliqué la coloration MGG sur le frottis pendant 3 minutes, puis on rincé la lame de sang, après cette opération on a appliqué la coloration Giemsa pendant 15 minutes et on a refait le rinçage, après cette dernière application on a laissé la lame pour sécher.[4]

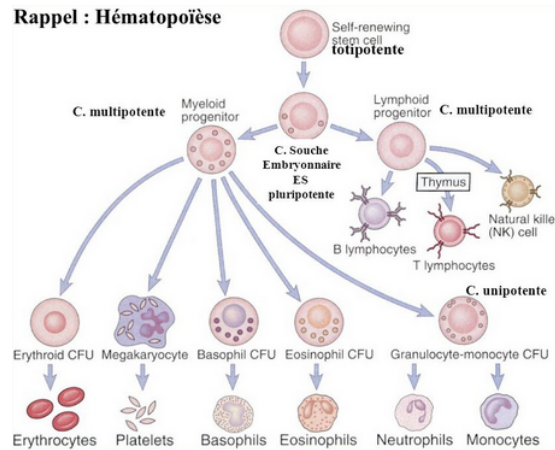


FIGURE 1.4 – Hématopoïèse

[3]

1.5 Les éléments figurés du sang périphérique

Le sang est composé de cellules sanguines en suspension dans le plasma. Les éléments figurés du sang ont des durées de vie limitées ; il existe un équilibre dynamique entre leur production (l'hématopoïèse et la lymphopoïèse) et leur destruction. L'hématopoïèse est la production des précurseurs sanguins (prolifération, différenciation et maturation) et se déroule dans les organes hématopoïétiques (moelle osseuse chez l'adulte, foie et rate chez l'embryon). [6]

1.5.1 Les globules rouges (Erythrocytes ou Hématies)

Les globules rouges (GR) sur frottis ont une forme circulaire, une taille uniforme avec un diamètre moyen de 8 nanomètre (de très petites variations de taille ou de forme sont observées normalement), 5 tera / l (millions par mm³), une coloration gris rosé avec un centre plus clair qui se fond graduellement à un anneau périphérique plus coloré. [6]

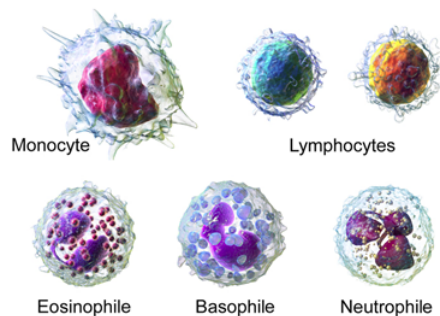


FIGURE 1.5 – Les types de globules blancs.

[6]

1.5.2 Les globules blancs (ou leucocytes)

Le préfixe leuc(o)- vient du grec leukos qui signifie « blanc » et l'élément -cyto vient du grec kutos qui signifie « cellule ». Les leucocytes jouent un rôle très important dans les défenses immunitaires. Il y a de 4.000 à 11.000 leucocytes dans un microlitre, ils se répartissent en :

- a. Polynucléaires ou granulocytes, impliqués dans les infections bactériennes, parasitaires et dans les allergies (40 à 80 % des leucocytes)
- b. Les mononucléaires qui comprennent en fait deux types de cellules totalement différentes :
 - Les monocytes qui ont une fonction de phagocytose (2 à 10% des leucocytes)
 - Les lymphocytes qui sont impliqués dans la reconnaissance spécifique des antigènes (20 à 40 % des leucocytes).

A. Les monocytes

Ces cellules ont une durée de vie dans le milieu sanguin très courte (environ 24 heures). En microscopie optique, elles apparaissent arrondies, ayant un diamètre de 15 à 20 nanomètres.

B. Les lymphocytes

Ce sont des cellules mononucléées, au rapport nucléo / cytoplasmique élevé. Leur durée de vie est variable, certains lymphocytes mémoires peuvent avoir une durée

de vie très longue.

En microscopie optique, ce sont des cellules de petites tailles, environ 7 nanomètres de diamètre avec un noyau occupant la quasi totalité de la cellule. Leur forme est régulière et arrondie. Il existe une petite frange cytoplasmique périphérique d'aspect mauve au MGG. Le noyau est sphérique, dense.

C. Les polynucléaires

Ce groupe de cellules possède des caractéristiques communes. Elles contiennent un noyau plurilobé. Les lobes sont reliés les uns aux autres par des ponts fins de chromatine.

Le cytoplasme est gris bleuté (ciel d'orage) au MGG et a un aspect un peu granuleux. Il existe en périphérie des voiles cytoplasmiques, visibles en microscopie optique. Le noyau est central, en fer à cheval ou en E.

1. Polynucléaires neutrophiles

Ce sont des cellules arrondies de 12 à 14 nanomètres de diamètre, caractérisées par la forme lurilobée de leur noyau (3 à 5 lobes). Les lobes sont réunis par de fins filaments de substance nucléaire qui ne sont pas toujours visibles si les lobes sont partiellement superposés.

La chromatine est dense, formée de masses sombres séparées par de petites bandes plus claires. Le cytoplasme contient d'assez nombreuses granulations assez fines, beige rosé, qui correspondent aux granulations spécifiques neutrophiles. Les granulations azurophiles ont, à ce stade de maturation, modifié leur affinité tinctoriale et ne sont presque plus visibles en optique .

2. Polynucléaires éosinophiles

Ce sont des cellules de 12 à 14 nanomètres de diamètre caractérisées par un noyau bilobé et surtout par l'aspect des granulations qui sont sphériques (0,5 à 1,5 nm de diamètre), réfringentes, de coloration orangée et assez nombreuses à l'intérieur du cytoplasme.

Leur durée de séjour dans le sang est 12 à 24 heures avant leur passage tissulaire et le temps de transit médullaire est comparable à celui du polynucléaire

neutrophile.

3. Polynucléaires basophiles

Ce sont des cellules de 10 à 14 nm de diamètre. Leur noyau bilobé est masqué par des granulations spécifiques qui sont assez nombreuses et dispersées sur toute la cellule. Elles sont arrondies (0,2 à 1 nm de diamètre) ou plus souvent angulaires, de coloration pourpre .

1.5.3 Les plaquettes ou Thrombocytes

Formées dans la moëlle osseuse à partir du mégacaryocyte, ce sont des structures discoïdes, anucléées (150 à 400 G/l). Leur durée de vie est de 8 à 10 jours.

- Les granules alpha contiennent de nombreuses protéines.
- Les granules denses contiennent de l'ADP, du calcium et de la sérotonine. [6]

1.6 Conclusion

Malgré les performances actuelles des automates qui identifient les populations normales et signalent les caractéristiques de cellules non identifiées, la reconnaissance des cellules pathologiques au microscope reste une étape indispensable pour un rendu optimal de l'hémogramme.[6]

Dans le chapitre suivant, on va présenter des généralités sur l'image et ses caractéristiques, ainsi que les espaces couleurs.

Chapitre 2

Traitement d'images et espace couleur

2.1 Introduction

Le traitement d'image est, avant tout, une science et une technique par essence multidisciplinaires. De par les connaissances qu'il est nécessaire de maîtriser sur la nature du signal représenté, le développement d'un outil de traitement d'image ne se conçoit pas en l'absence de solides notions sur la discipline au sein de laquelle l'image est exploitée. Ces connaissances vont de l'art pictural aux mathématiques, en passant par les sciences humaines, la médecine et les sciences expérimentales.

L'opération de formation d'une image conditionne elle aussi fortement ses caractéristiques et une connaissance approfondie de la physique qui sous-tend la création du signal et sa détection est elle aussi de première importance.[7]

Dans ce chapitre nous présenterons tout d'abord ce que c'est l'image et sa texture sans oublier sa représentation dans les espaces couleurs

2.2 Le pixel

Le pixel (picture element) est la taille du plus petit élément de l'image. Il désigne aussi un point de la matrice, il n'a pas de dimension, pas de aire. Le pixel peut avoir

une dimension.

Un pixel possède une valeur qui peut être un scalaire et représenter un niveau de gris ou un vecteur représentant une couleur, ou tout autre chose.

Les images dites en "noir et blanc" sont composées de pixels binaires noirs ou blancs [deux valeurs possibles). Les images en niveaux de sont composées de pixels de valeurs scalaires Représentant la luminosité. Pour donner un ordre de grandeur, si un pixel est codé sur 8 bits (1 octet), on dispose de $2^8 = 256$ couleurs ou niveaux de gris. Sous Matlab, le pixel (1,1) est située en haut à gauche de l'image.

En général on utilise plutôt le pixel physique (liée à l'écran). Ainsi un pixel physique peut représenter plusieurs points sur l'image réelle ou inversement[8].

2.3 Généralités sur l'image

2.3.1 Définition d'une image

Une image est une représentation planaire d'une scène ou d'un objet situé en général dans un espace tridimensionnel. Elle peut être décrite sous la forme d'une fonction $I(x, y)$ tel que : x et y sont des coordonnées spatiales d'un point de l'image et I est une fonction d'intensité lumineuse (niveau de gris et de couleur). Cette valeur est généralement proportionnelle à l'énergie rayonnante reçue dans la bande de fréquence électromagnétique. Sous cet aspect, l'image est inexploitable par la machine, ce qui nécessite sa numérisation[9]

Les informations d'une image sont :

- nombre de lignes,
- nombre de colonnes,
- format des pixels (bits, niveaux de gris, niveaux de couleurs),
- compression éventuelle[10]

2.3.2 Typologie d'image

Les formats d'images sont à diviser en deux grandes familles :

A. Les images (matricielles) ou bitmap

Une image matricielle (ou bitmap) est une image constituée d'un ensemble de points : les pixels. Chaque point porte des informations de position et de couleur. Format d'images bitmap : BMP, PCX, GIF, JPEG, TIFF. Les photos numériques et les images scannées sont de ce type sont composées de formes géométriques qui vont pouvoir être décrites d'un point de vue mathématique. Par exemple une droite sera définie par 2 points, un cercle par un centre et un.

B. Les images vectorielles

Les images vectorielles rayon. Le processeur est chargé de "traduire" ces formes en informations interprétables par la carte graphique (images Word, Publisher, CorelDraw - format WMF, CGM, etc.)[8]

Dans notre cas, on s'intéresse à l'image couleur qui est avant tout une image numérique

2.3.3 Texture de l'image

En traitement d'image, de nombreux auteurs et chercheurs ont essayé de donner une définition scientifique satisfaisante et universelle de la texture sans toutefois n'y être jamais parvenues.

Le glossaire standard de l'IEEE de terminologie de traitement d'image et reconnaissance de formes en donne une définition très générale, "la texture est un attribut décrivant l'arrangement spatial des niveaux de gris des pixels d'une région", Certaines caractéristiques texturales ont été reconnues par plusieurs chercheurs (Tamura et al., 1978 ; Laws, 1980b) comme jouant un rôle fondamental dans la discrimination par le cerveau de différentes textures. Parmi ceux-ci, on peut citer : l'uniformité, la densité, la granularité, la rugosité, la régularité, la linéarité, la directionnalité, la direction, la

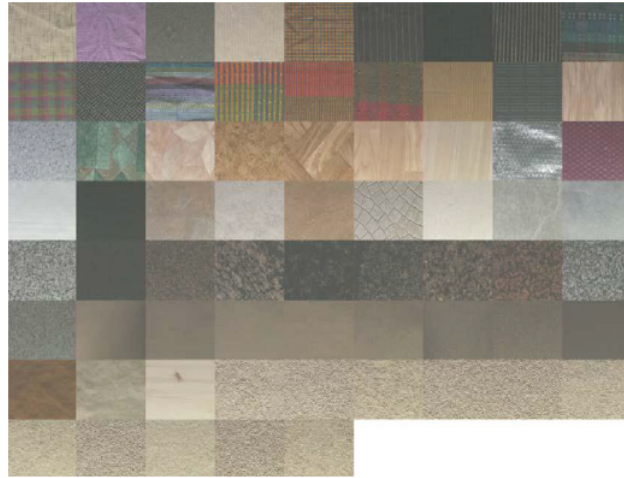


FIGURE 2.1 – Image de texture couleur de la base Outex
[11]

fréquence et la phase. Comme on peut le voir, Ces dernières ne sont pas indépendantes et peuvent être classées dans trois grandes catégories : la direction, la fréquence et le chaos[10]

2.3.4 Technologie de l'image numérique

1. image binaire : $p = 1$ et $M = 1$
2. image en niveau de gris : $p = 1$ et $M = 255$
3. image couleur : $p = 3$ et $M = 255$

A. Echantillonnage de l'image continue

Sachant que la représentation informatique d'une image est nécessairement discrète alors que l'image est de nature continue : le monde est continu.

Certains capteurs effectuent une discrétisation, c'est le cas des appareils photo nu-

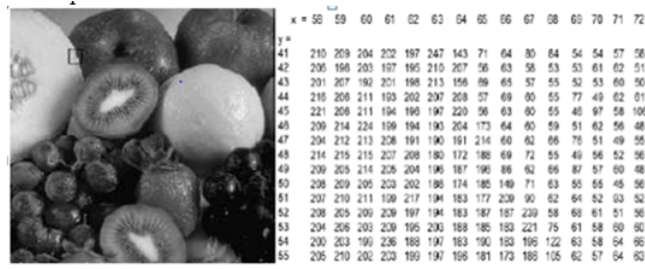


FIGURE 2.2 – Image numérique

[7]

mériques, des scanners. Si on regarde d'un peu plus près, la transformation d'un signal analogique 2D nécessite à la fois une discrétisation de l'espace, c'est l'échantillonnage, et une discrétisation des couleurs, c'est la quantification. L'image numérique est obtenue par échantillonnage (elle doit respecter le théorème de Shannon). En pratique, on s'appuie sur conversion analogique/numérique et réciproquement grâce aux différents filtres.

B. La quantification

Une fois l'image échantillonnée, l'étape suivante de la numérisation est celle de la quantification. Le débit binaire y est directement lié.

- Pour quantifier un signal, il faut définir une loi et une échelle de quantification.
- La quantification a été établie sur 256 niveaux de gris, soit 8 bits, et la loi a été choisie linéaire.
- Pour le noir et blanc, 2 bits de quantification sont suffisants ; pour la couleur à laquelle l'oeil est moins sensible moins 24 bits de quantification suffisent, d'où le choix de travailler sur des octets.
- Cela satisfait à la fois les contraintes de la luminance et celles de la chrominance.

2.3.5 Image noir et blanc

L'image noir et blanc est une image binaire dont tous ses pixels sont soit de valeur 1 ou 0, la première valeur définie la couleur blanche qui simule tous les objets existants dans l'image, et le zero sera attribué au fond de l'image, ce genre d'images prend le

moins d'espace mémoire par rapport aux autres.

2.3.6 Image au niveau de gris

On passe au niveaux de gris en considérant comme seule composante la luminosité. Il existe plusieurs manières de convertir une image RVB en niveaux de gris. La plus simple est de faire

$$gris = rouge + vert + bleu$$

C'est aussi équivalent affecter la couleur gris à chacune des trois composantes RVB. L'idéal est de faire ressortir la luminosité d'un pixel. Celle-ci vient principalement de la présence de la couleur verte. On emploie généralement les coefficients suivant

$$gris = 0,299 \cdot rouge + 0,587 \cdot vert + 0,114 \cdot bleu$$

2.3.7 Résolution d'une image

La résolution d'une image est définie par le nombre de pixels par unité de longueur de la structure à numériser i.e classiquement en dpi *dots per inches* ou ppp points par pouce. Les phénomènes de numérisation dépendent des 2 équations suivantes : $(X \cdot \text{résolution}) = x \text{ pixels}$ $(Y \cdot \text{résolution}) = y \text{ pixels}$ où X et Y représente la taille géométrique (en pouces ou centimètres) de la structure à numériser, résolution représente la résolution de numérisation, puis x et y représente la taille (en pixels) de l'image. [12] Une image de 100 ppp (ou 100 dpi) contient 10 000 points par pouce carré. $(100 \times 100 = 10\,000)$ La résolution définit la netteté d'une image et sa qualité d'affichage à l'écran. Plus la résolution est grande (c'est-à-dire plus il y a de pixels dans une longueur de 1 pouce), plus votre image est précise dans les détails. C'est un critère important lorsque l'on utilise un scanner. [9]

2.3.8 Taille d'une image

Une image de 1 pouce *1 pouce scannée à 100 dpi aura une taille (x, y) de 100 pixels sur 100 pixels c'est à dire $(1*100)*(1*100) = 100$ pixels sur 100 pixels = 10.000 pixels² ou simplement 10.000 pixels .

Remarque :

1 pouce = 2,54 cm

1 pouce = 25,40 mm = 100 pixels (dans le cas de l'exemple actuel choisi)

1 inch = 2,54 cm = 1 pouce

1 f t = 12 inch

2.4 Définition des grandeurs chromatiques

- **La teinte** : « est un attribut de la sensation visuelle qui peut être décrit par des qualificatifs tels rouge, jaune... [12]» la teinte détermine donc la couleur d'une surface.
- **Le chroma** : « est le niveau de coloration d'une surface, évalué relativement à la lumière qu'elle reçoit. Le chroma d'une surface donnée est un attribut perceptif indépendant du niveau d'éclairement. Pour une surface de chromaticité constante, le chroma augmente avec la clarté (ou la luminosité) de la surface, contrairement à la saturation [12]»
- **La saturation** : « est une grandeur permettant d'estimer le niveau de coloration d'une teinte indépendamment de la luminosité. Elle représente la pureté de la couleur perçue comme vive, pâle, terne,... [14]».
- **La luminosité** : « est un attribut d'une sensation visuelle selon laquelle une surface paraît émettre plus ou moins de lumière. traduite par des vocables comme clair, foncé, lumineux, beaucoup de grandeurs physiques permettent de la quantifier. Ainsi les termes intensité, éclairement, luminance, clarté, sont souvent employés dans la littérature pour désigner la notion de luminosité[13].» Les images que nous

utiliserons dans notre étude sont de type couleur.

2.5 Exploitation de la couleur dans une image couleur

Chaque image numérique couleur I qui est une matrice de pixels P , dont la couleur $C(P)$ est définie par trois composantes rouge, verte et bleue notées respectivement $CR(P)$, $CG(P)$ et $CB(P)$. La couleur peut non seulement être représentée en utilisant ces trois composantes, mais aussi faire appel à d'autres systèmes de représentation de la couleur. Une étude détaillée de ces différents systèmes de représentation est faite dans la suite de ce chapitre. De manière générale, à partir de l'image couleur I codée dans l'espace $(\vec{C}_1, \vec{C}_2 \text{ et } \vec{C}_3)$ dont C_1 , C_2 et C_3 sont les composantes de n'importe quel système utilisé, on peut extraire trois image composantes notées ICK en tenant compte à chaque fois d'une composante primaire parmi les trois utilisées et dont chaque pixel est caractérisé par une des composantes couleur C_k

$$k \in (1, 2, 3)$$

.Nous appellerons d'autre part IL , l'image de luminance correspondant à l'image couleur I (l'image de luminance veut dire monochromatique). Dans le cas général, la couleur d'un pixel est représentée par trois composantes notées C_1 , C_2 et C_3 . A ces trois composantes, correspondent respectivement trois vecteurs directeurs normés qui forment le repère d'un espace vectoriel d'origine O , appelé espace couleur et noté $(\vec{C}_1, \vec{C}_2 \text{ et } \vec{C}_3)$. Dans cet espace, la couleur d'un pixel donne naissance à un point C qui définit le vecteur couleur OC et dont les coordonnées sont les niveaux des composantes C_1 , C_2 et C_3 .

2.6 Espaces colorimétriques

D'après le principe de la trivariance visuelle : toute couleur peut être reproduite visuellement à l'identique dans des conditions d'observation déterminées par le mélange algébrique en proportions définies de manière unique, de trois couleurs, appelées primaires. Ces trois couleurs peuvent être arbitrairement choisies sous la seule réserve

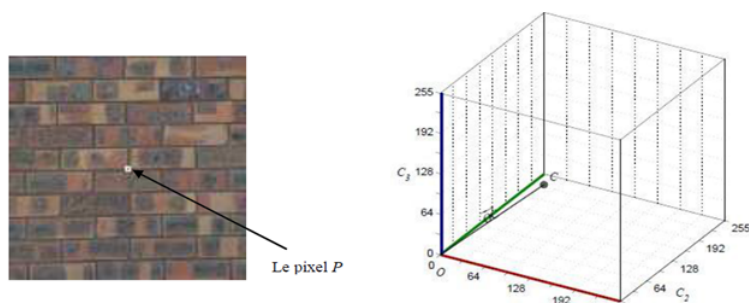


FIGURE 2.3 – Le pixel encadré en blanc de composantes couleur sont C_1 , C_2 et C_3 donne naissance au point C

[11]

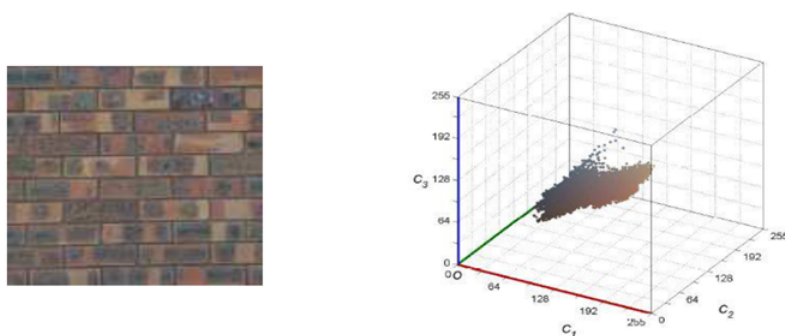


FIGURE 2.4 – Les couleurs des pixels des images donnent naissance aux nuages de points.

[11]

qu'aucune d'entre elles ne puisse être reproduite par un mélange approprié des deux autres. Ce principe est à l'origine de tous les espaces de représentation de la couleur qui sont nombreux. Chacun présente des caractéristiques particulières. Ils peuvent être regroupés en quatre familles principales.

1. Les espaces de primaires.
2. Les espaces de luminance-chrominance.
3. Les espaces perceptuels.
4. Les espaces d'axes indépendants.

Dans la suite de ce chapitre, on rencontrera souvent la notion de blanc de référence et la notion d'illuminant. En effet, il est important de les définir car ces deux notions rentrent dans la constitution de chaque espace. Le blanc de référence permet de fixer les valeurs unitaires des trois primaires utilisées et pour le modéliser, la CIE (Commission Internationale d'Eclairage) a fait des normalisations vis-à-vis de la notion d'illuminant qui représente la répartition spectrale relative d'énergie des différentes lumières émises par le soleil ou par n'importe quelle source de lumière.[14]

2.6.1 Les espaces de primaires

En 1931, la CIE a défini un ensemble de trois couleurs primaires, le rouge noté [RC], le vert noté [GC] et le bleu noté [BC]. Ces trois primaires ont été choisies de telle sorte que le mélange unitaire reproduise l'impression du spectre équi-énergétique (illuminant E).

A. L'espace RGB

L'espace RGB (Red Green Blue) ou bien RVB(en version française) est un système additif défini à partir des trois primaires monochromatiques de couleur rouge, verte et bleue. Avec le principe que chaque stimulus de couleur est représenté par un point C qui définit le vecteur couleur OC. Les coordonnées de ce vecteur sont les composantes trichromatiques RC, GC et BC. Les points correspondants à des stimuli de couleur sont contenus dans un cube, connu sous le nom de cube des

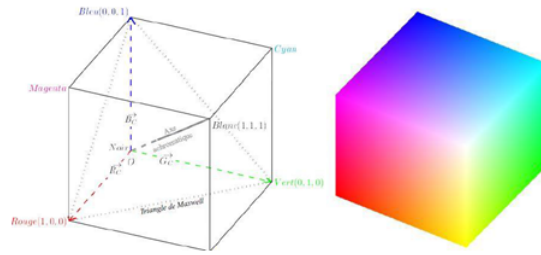


FIGURE 2.5 – Le cube des couleurs.

[11]

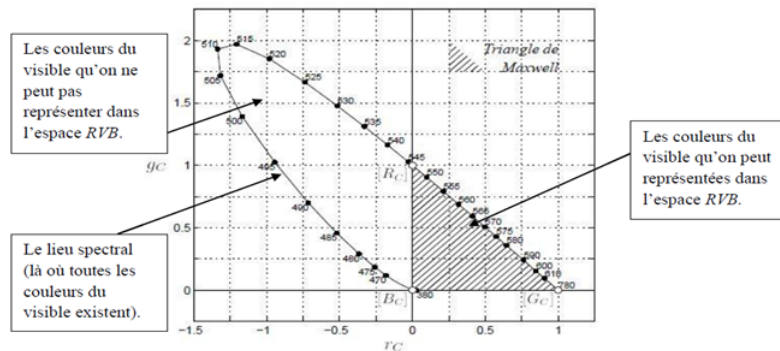


FIGURE 2.6 – Le diagramme de chromaticité (r_c , g_c) de la CIE

[11]

couleurs (figure 2.5). L'origine O correspond au noir

$$(R_c = G_c = B_c = 0)$$

tandis que le blanc de référence est défini par le mélange unitaire des trois primaires

$$(R_c = G_c = B_c = 1)$$

. La droite passant par les points Noir et Blanc est appelée axe des gris, axe des couleurs neutres ou encore axe achromatique. En effet, les points de cette droite représentent des nuances de gris allant du noir au blanc. La définition de l'espace de représentation de la couleur, n'étant pas unique, puisqu'il dépend des primaires ainsi que du blanc de référence. Pour pallier les inconvénients des systèmes RVB. La CIE (Commission International d'Eclairage) a défini un autre espace nommé XYZ, qui permet de représenter toutes les couleurs (Figure 2.8).

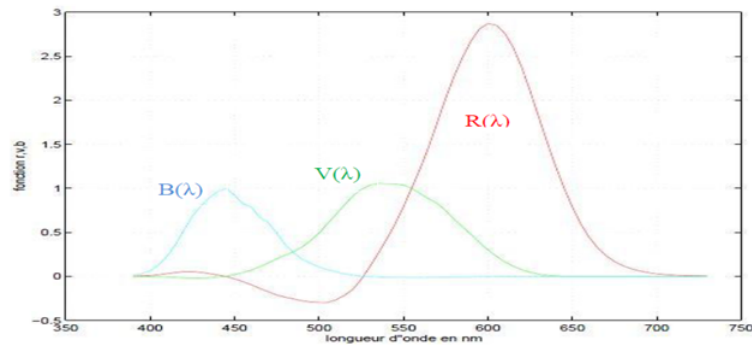


FIGURE 2.7 – Représentation des fonctions colorimétriques R,V,B

[7]

B. L'espace XYZ

En 1931, les travaux de Judd ont permis à la CIE (Commission International d'Eclairage) d'établir le système de référence colorimétrique XYZ dont les primaires (X, Y, Z) sont virtuelles (ou imaginaires) et la composante Y représente la luminance. En effet, il correspond à un changement de primaires qui s'obtiennent facilement à l'aide d'une matrice de passage à partir de n'importe quel système RVB. Par exemple, le passage du système RVB de la CIE sous l'illuminant E qui correspond à une lumière d'énergie constante (ou blanc équiénergétique) au système XYZ s'effectue comme suit :

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 2,7690 & 1,7518 & 1,1300 \\ 1,000 & 4,5907 & 0,0601 \\ 0,000 & 0,0565 & 5,5943 \end{pmatrix} \begin{pmatrix} R \\ V \\ B \end{pmatrix}$$

Dans tous les cas, à partir de ce résonnement, toute les conversions entres les différents systèmes colorimétriques sont possible, et cela est valable même pour les autres systèmes qu'on n'a pas encore vue.

2.6.2 Les espaces de luminance-chrominance

Ce type d'espace possède une composante appelée luminance, qui permet de quantifier la luminosité, et deux autres composantes de chrominance notées (chr1 et chr2)

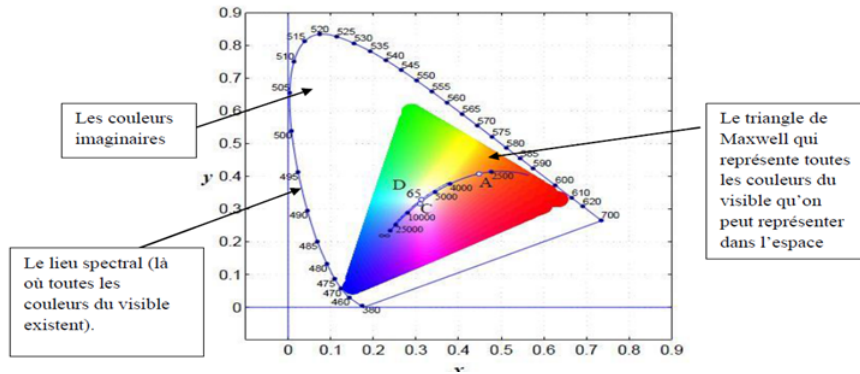


FIGURE 2.8 – Le diagramme de chromaticité (x, y) de la CIE.

[15]

qui sont suffisantes pour quantifier le caractère chromatique d'un stimulus de couleur. Il existe différents espaces luminance-chrominance. Cette famille est divisée par Vandembroucke en quatre sous familles qui sont intitulées :

- Les espaces perceptuellement uniformes (où la luminance est notée L).
- Les espaces antagonistes (où la luminance est notée A).
- Les espaces de télévision (où la luminance est notée Y).
- Les autres espaces luminance-chrominance.

2.6.3 Les espaces perceptuels

Selon Vandembroucke et plusieurs autres auteurs, disent que l'œil humain ne perçoit pas la couleur comme une combinaison de composantes trichromatiques (R, V, B) mais selon des entités plus subjectives liées à la luminosité L (en anglais *Lightness* ou *brightness*), la teinte H (en anglais *Hue*) et la saturation S (en anglais *Saturation*). Donc les espaces couleur perceptuels décrivent la couleur à partir de ces trois composantes Luminosité, teinte et saturation dont leur appellation générique est LTS.

Il existe de nombreux systèmes de ce type dans la littérature, présentés sous différentes dénominations telles que HSI, HSV, HSL, TLS, LCH, LSH, LST, ITS, . . . , ces différentes dénominations désignent les mêmes grandeurs mais se différencient par leur calcul.

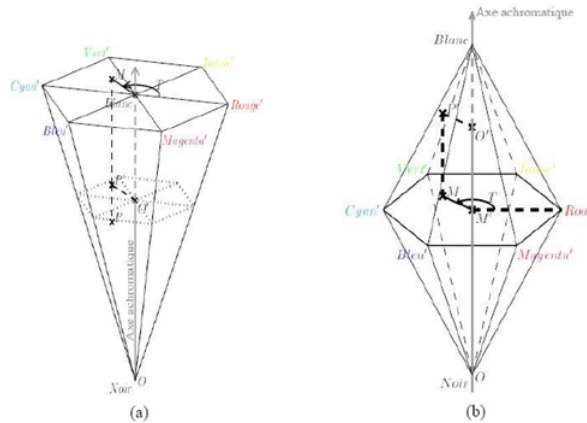


FIGURE 2.9 – Le système HSV de cône hexagonal simple (a) et double (b).

[15]

En raison de multitude de ce type d'espace, On ne va pas les étudier tous, on a choisi seulement de faire les deux premiers c'est-à-dire HSV et HSI car ils sont les plus utilisés en traitement d'images[14]

A. L'espace HSV (Hue, Saturation, Value)

Le système HSV est connu aussi sous le nom de système de cône hexagonal. Il présente deux variante : le modèle de cône hexagonal simple et le modèle hexagonal double, comme l'indique la figure (Figure 2.10).

Toutes ces expressions sont calculées à l'aide des règles géométriques et trigonométriques qu'on n'a pas détaillées. L'utilisation de la nomination (R, G, B) pour définir le système de primaire RVB revient à ne pas confondre le V du HSV avec le V du RVB.

B. L'espace HSI (Hue, Saturation, Intensity)

Cet espace est connu aussi sous le nom de coordonnées triangulaires, comme son nom l'indique, il est basé sur une forme géométrique particulière aux surfaces d'iso-intensité. C'est à l'aide du cube des couleurs RVB qu'on a vu précédemment que la forme géométrique est formée. Le cube est basculé autour du sommet O représentant le noir, de façon que la diagonale achromatique prenne une position

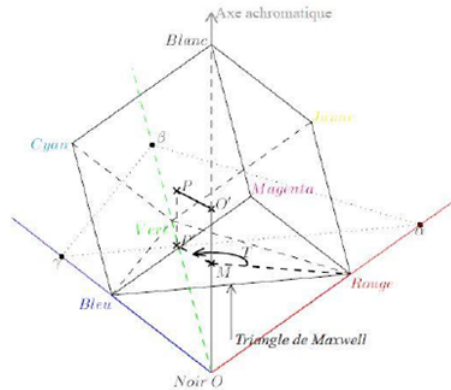


FIGURE 2.10 – Le cube des couleurs dans le modèle triangulaire.

[11]

verticale (Fig. 10). Maintenant soit P le point représentatif d'une couleur donnée. L'intensité I est fournie par la projection de P sur l'axe achromatique vertical, la teinte H est repérée par l'angle que fait la droite OP avec l'axe achromatique, et la saturation S est égale à la distance de P à l'axe achromatique.[16]

2.6.4 Les espaces d'axes indépendants

Toutes les composantes des espaces qu'on a vu jusqu'ici sont plus ou moins corrélées (plus pour les espaces de primaires, moins pour les espaces perceptuels et perceptuellement uniformes qui décorrèlent la composante de luminance des composantes de chrominance). Le but de ce type d'espace est d'établir une décorrélation complète entre les composantes de couleur, c'est-à-dire toutes les composantes seraient indépendantes [15].

Il existe deux techniques pour remédier à ce problème de corrélation entre les composantes couleur, l'une est l'analyse en composantes principales (ACP), l'autre est l'analyse en composantes indépendantes (ACI). Une décorrélation complète peut être obtenue par l'utilisation de la technique d'analyse en composantes principales (ACP)[13].

2.7 Conclusion

Dans ce chapitre, nous avons donné un aperçu sur l'image et ses caractéristiques, ainsi que nous avons abordés des généralités sur que les principaux espaces couleurs. Dans le chapitre suivant, nous nous intéresserons à l'hématologie cellulaire en définissant quelques termes du domaine qui présentent notre point d'intérêt.

Chapitre 3

Etat de l'art sur la segmentation d'images

3.1 Introduction

L'analyse d'images a pour but l'extraction de l'information caractéristique contenue dans une image. Le résultat d'une telle analyse s'appelle très souvent la description structurelle. Celle-ci peut prendre la forme d'une image ou de toute structure de données permettant une description des entités contenues dans l'image. Par opposition avec la phase d'interprétation, l'analyse tente, dans la mesure du possible, de ne pas prendre en compte le contexte (i.e. l'application). Essentiellement, l'analyse de l'image fait appel à la segmentation où l'on va tenter d'associer à chaque pixel de l'image un label en s'appuyant sur l'information portée (niveaux de gris ou couleur), sa distribution spatiale sur le support image, des modèles simples (le plus souvent des modèles géométriques). La segmentation d'images ainsi définie est un domaine vaste où l'on retrouve de très nombreuses approches. [17]

Dans ce chapitre, on va consacrer une première partie pour donner un aperçu sur la classification et la segmentation d'images, ensuite, nous nous intéresserons aux différentes approches de la segmentation d'images et à la fin on définit une des méthodes de seuillage qu'on a utilisé dans notre travail.

3.2 Problématique classification-segmentation

Le domaine qui regroupe l'ensemble des méthodes et techniques opérant sur les images est le traitement et analyse d'images. Ce dernier s'intéresse, entre autres, à l'amélioration de l'aspect visuel de l'image (filtrage, restauration,...) et à l'extraction des informations jugées pertinentes (Extraction de primitives et segmentation). En analyse d'images [18], il est important de bien distinguer classification et segmentation, même si ces deux notions sont très proches.

3.2.1 Classification

La classification est, elle, une méthodologie consistant à classer des individus (des pixels, des images,...) en exploitant leur similarité, mais en aucun cas on ne cherche à partitionner l'image. Cependant, la classification de pixels peut être une étape de la segmentation d'images. [19]

La classification consiste à attribuer l'échantillon à l'une des classes (dans associer l'échantillon à l'un des prototypes). La tâche de classification est effectuée de la manière suivante : chaque échantillon de l'ensemble de données est comparée à l'ensemble des expressions et affecté à l'un d'eux (dire à une classe ou sous-classe) ou rejetée.

3.2.2 Segmentation

La segmentation est un traitement qui consiste à créer une partition de l'image considérée, en sous-ensembles appelés régions. Une région est un ensemble connexe de pixels ayant des propriétés communes (intensité, texture,...) qui les différencient des pixels des régions voisines.[19]

Zucker définit la segmentation d'image comme le partitionnement de l'ensemble des pixels d'une image en sous-ensembles appelées régions R_i , avec i allant de 1 à n , et I de R_1 à R_n . telle que aucune région ne doit être vide, l'intersection entre deux régions

doit être vide et l'ensemble des régions doit recouvrir toute l'image. Une région est un ensemble de pixels connexes ayant des propriétés communes qui les différencient des pixels des régions voisines.

Cette définition se traduit mathématiquement par les relations suivantes :[16]

$$\bigcup_{i=1}^n R_i = I \quad \text{avec} \quad R_i \cap R_j = \emptyset \quad i \neq j$$

$$\text{et} \quad \begin{cases} P(R_i) = \text{vrai} & \forall i = 1, \dots, n \\ P(R_i \cup R_j) = \text{faux} & R_i \text{ adjacente à } R_j \end{cases}$$

3.3 Différentes approches de la segmentation d'images

Il existe une multitude de méthodes de segmentation qu'on peut regrouper en deux grandes catégories [20]

- Segmentation fondée sur les contours (approche contour).
- Segmentation fondée sur les régions (approche région).

3.3.1 Approche contour

Cette catégorie correspond à une approche duale qui cherche à détecter une transition (ou contour) entre deux régions connexes. Les détecteurs de contours peuvent être simples, comme les opérateurs de Sobel ou de Roberts, ou plus complexes tel que l'opérateur de Canny. La segmentation par approche contour peut être faite à partir de l'image de gradient, du laplacien par détection des contours ou par application d'un contour déformable initialisé automatiquement ou manuellement.

3.3.2 Approche région

Nous retrouvons dans cette catégorie la segmentation basée sur la croissance de régions et les méthodes par division-fusion. La croissance de régions débute à partir de germes dans l'image que l'on propage aux pixels voisins de teintes ressemblantes. Nous agglomérons ainsi les pixels de proche en proche, jusqu'à ce qu'aucun regroupement ne

soit plus possible. Les méthodes dites de division partent de l'image entière que nous irons la découper récursivement en plus petites régions tant que ces régions ne sont pas suffisamment homogènes. Une approche duale de fusion part d'un ensemble de petites régions homogènes qui sont récursivement regroupées sous des conditions de proximité spatiale (adjacence) et d'homogénéité. Les méthodes dites de division-fusion sont un mélange de ces deux approches. Les algorithmes que nous allons présenter et commenter sont :

1. Accroissement de région fonctionnant au niveau du pixel.
2. Split and merge classique (d'après [21]).
3. Algorithme CSC (d'après [22]) : un algorithme de « merge and split » qui donne d'excellents résultats et qui possède un certain nombre de propriétés intéressantes.

1. Croissance de région

Cet algorithme a été développé à l'INPE et mis en œuvre dans le géographique Information et Image Processing System- PRINTEMPS (DPI et al, 1995). Il a été intensivement utilisé dans la segmentation de la région amazonienne images pour évaluer la terre utilisé changements. Certaines segmentations résultantes de LANDSAT Images TM sont également présentées[23]. Dans le domaine médical, l'algorithme de croissance de région a été utilisé avec succès comme technique de segmentation d'images médicales et numériques (Ngah et al., 2002, Venkatachalam et al., 2002, Tuduki et al., 2000, Ooi et al., 2000, Justice & Stokely, 1997).

Diverses modalités d'imagerie médicale tels que les rayons X, la tomographie par ordinateur (CT) et l'imagerie par résonance magnétique (IRM) sont largement disponible et utilisé en routine clinique pratique. Il a été généralement accepté que l'interprétation visuelle de ces images est très subjective. En conséquence, il existe un grand besoin pour les techniques de traitement d'image qui doit être

appliquée sur les images médicales pour aider les pathologistes et les radiologues pour faciliter et avoir un meilleur diagnostic des maladies. L'une des techniques de traitement d'image est commune dispose d'extraction. Applications de divers techniques d'extraction des caractéristiques dans les images médicales ont augmenté ces dernières années (Christoyianni et al., 1999, Chen et al., 1998, Pham et al., 1997, El-Faramawy et al., 1996, Chitre et al., 1993).

L'étude courante est allée plus loin en utilisant l'utilisation potentielle de seuillage pour l'algorithme de croissance de région comme technique d'extraction des fonctionnalités. L'algorithme proposé d'extraction de caractéristiques est appelé « seeded region growing features extraction » (SRGFE). Cet algorithme a été utilisé pour extraire des caractéristiques des cellules ; taille du noyau, la taille de cytoplasme et niveau de gris du cytoplasme.

Le test de corrélation a été appliqué entre les données extraites en utilisant l'algorithme proposé SRGFE avec les données extraites manuellement par les cytologistes. La forte valeur de corrélation obtenue dans le test de corrélation montre que l'algorithme SRGFE est adapté et a une haute capacité à être utilisé comme une technique d'extraction d'image pour extraire les caractéristiques importantes des cellules cervicales. Cela aiderait les cytopathologistes et les cytotechnologistes dans le dépistage du cancer du col de l'utérus en fournissant La valeur précise de la taille et de niveau de gris caractéristiques nucléaires et cytoplasmiques[24].

2. Méthodes du type fusion et division (*split and merge*)

L'algorithme *split and merge* a été présenté la première fois en 1974 par Pavlidis et Horowitz [25]. Cet algorithme s'apparente dans son principe à l'algorithme d'accroissement de région que nous venons de présenter. La différence principale provient de la nature des régions élémentaires agrégées. Dans l'algorithme (*split and merge*), les régions agrégées proviennent d'une première phase (*split*) de traitement de l'image qui construit de manière récursive des régions carrées de taille

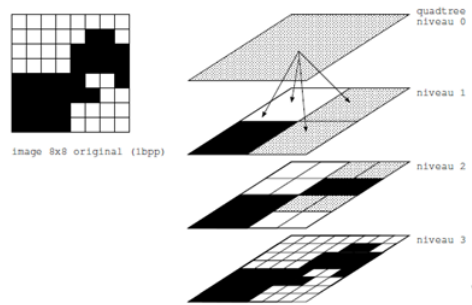


FIGURE 3.1 – Découpage par quadtree d'une image 8x8.

variable mais homogènes.

A. Split

La méthode de découpage de l'image utilisée dans cet algorithme est basée sur la notion de *quadtree*. Cette structure de données est un arbre quaternaire qui permet de stocker l'image à plusieurs niveaux de résolution. On part d'une région initiale qui est l'image tout entière. Si cette image vérifie un critère d'homogénéité de couleur, l'algorithme s'arrête. Sinon, on découpe cette région en quatre parties de même taille et on lance la procédure récursivement dans ces quatre parties. La région initiale va être stockée comme un nœud dans un graphe et les sous parties comme des fils de ce nœud. La figure 5 montre une image en noir et blanc 8x8 et le découpage correspondant à chaque niveau. Traitement avec seuil égale à 100 %

Dans cet exemple, le critère d'homogénéité est absolu : une zone est dite homogène si elle ne contient que des pixels de même couleur (seuil=100%). On peut être plus tolérant et accepter qu'une zone soit déclarée homogène dès que plus de 75% d'une couleur domine. La figure 7 montre le découpage correspondant. De manière plus générale, on va appliquer ce principe de réduction à des images colorées. Le critère d'homogénéité est fixé par un seuil sur la variance de la couleur dans la zone en cours d'examen. Au-dessus de ce seuil, la zone est découpée en quatre, en dessous, elle est conservée et constitue un

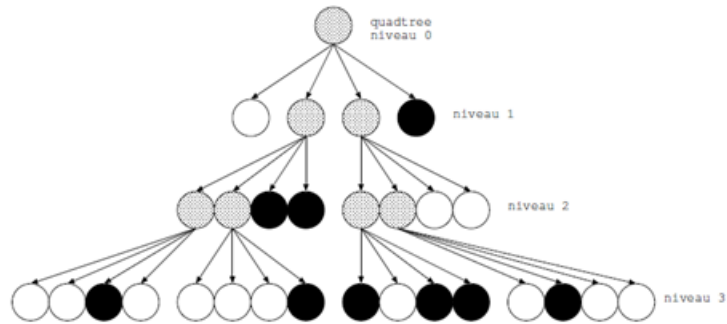


FIGURE 3.2 – Quadtree.

nœud terminal de l'arbre. On lui attribue alors la couleur de la moyenne des pixels la constituant.

B. Merge

La procédure de découpage décrite précédemment aboutit à un nombre de régions trop élevé. La cause fondamentale de cette sur-segmentation est que l'algorithme découpe les régions de manière arbitraire. Il se peut qu'il coupe de cette façon une zone homogène en deux ou quatre parties sur une image 8x8.

Des améliorations substantielles à cet algorithme existent (entre autres, [[20],[26], [27], [28]]) et nous n'avons présenté ici que la version la plus élémentaire du *split and Merge* .

C. **Limitations** Dans la figure ci-dessous, les parties 1, 2, 3 et 4 de cette image par exemple, appartiennent à des branches différentes du quadtree. Elles sont donc considérées comme des régions différentes bien que leur couleur soit identique.

La solution, qui correspond à la phase « merge » de l'algorithme, est de procéder à une fusion de régions après le découpage. L'implémentation la plus simple de cette fusion cherche tous les couples de régions adjacentes dans l'arbre issu du découpage et cherche à les fusionner si leur couleur est assez proche.

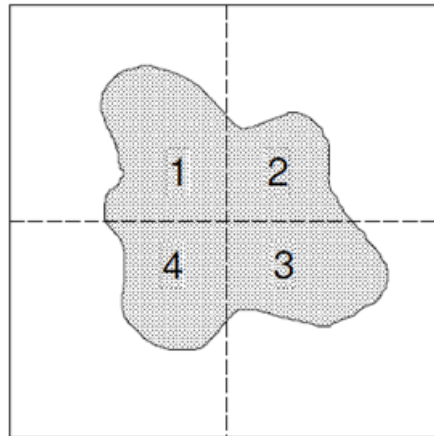


FIGURE 3.3 – Problème de découpage arbitraire de régions dans la phase de division.

3. **Algorithme CSC** L'algorithme CSC que nous présentons à présent combine beaucoup des avantages des méthodes précédentes mais ne présente pas les limitations que nous avons soulignées. L'algorithme CSC (*Color Structure Code*) a été présenté en 1993 par Rehrmann et Prieese. Il repose sur l'utilisation d'une structure hexagonale hiérarchique de codage de l'image. Il s'agit d'un algorithme de « merge and split » Particulièrement efficace et robuste. L'une des propriétés de cet algorithme est qu'il est complètement parallélisable et qu'il ne dépend pas d'un ordre de traitement de l'image comme les algorithmes d'accroissement de région.[29]

4. **La ligne de partage des eaux** Considère une image comme un relief topographique, dont on simule l'inondation. Cette approche suppose qu'on représente les images en termes géographiques, où l'on associe à chaque pixel une altitude relative à son niveau de gris(couleur ou texture). Il est alors possible de définir la ligne de partage des eaux (watershed) comme étant la crête formant la limite entre deux bassins versants.

Pour l'obtenir, il faut imaginer l'immersion d'un relief dans de l'eau, en précisant que l'eau ne peut pénétrer dans les vallées que par ses minima (voir figure 3.5). La ligne de partage des eaux est représentée par les points où deux lacs disjoints se rejoignent au cours de l'immersion. Elle possède l'avantage de

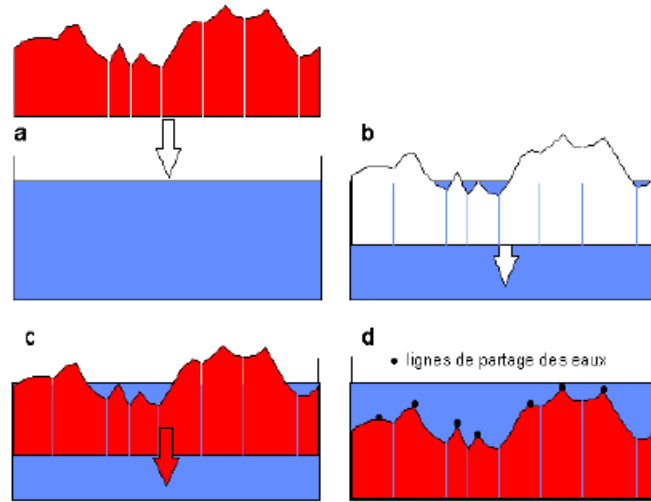


FIGURE 3.4 – Principe de segmentation watershed.

[30]

fournir des régions délimitées par des contours fermés formant une partition de l'image, mais présente l'inconvénient d'observer une sur-segmentation de l'image, faisant apparaître des lignes de crête en dehors des objets, ainsi qu'à l'intérieur d'objets convexes.

3.4 Segmentation basée sur les pixels

Le principe consiste à regrouper les pixels selon leurs attributs sans tenir compte de leur localisation au sein de l'image. Cela permet de construire des classes de pixels ; les pixels adjacents, appartenant à une même classe, forment alors des régions. Parmi les approches pixellaires nous distinguons les méthodes de seuillage et les méthodes de classification (clustering).

La classification consiste à regrouper et à classer les pixels d'une image en classes en fonction de leurs propriétés. Il existe deux grandes tendances : la classification supervisée (PMC, SVM, K-plus proches voisins,...), basée sur l'apprentissage de propriétés discriminantes sur un échantillon de données déjà classées, et la classification non su-

pervisée (K-means, GMM, Estimation-Maximisation(EM),...) basée sur une mesure de distance entre les vecteurs d'attributs. [19]

3.5 Seuillage global

3.5.1 Les méthodes non paramétriques

Ces méthodes permettent de trouver le seuil optimal de segmentation sans aucune estimation de paramètres. Généralement, ces méthodes sont basées sur l'optimisation de critères statistiques. Ces méthodes consistent à déterminer le seuil optimal à partir de l'histogramme de l'image. La méthode la plus connue est sans doute la méthode d'Otsu.[31]

D'autres méthodes de seuillage sont basées sur l'entropie de l'histogramme. On parle alors de seuillage entropique. Parmi ces méthodes, on peut citer les méthodes de Pun [32], Kapur [33], Johansen et Bille [34], de cross entropie [35], d'entropie de Renyi [36], etc.

– La méthode Otsu

Dans son papier original, Otsu décrit trois critères discriminants possibles : la variance intra-classe, la variance interclasse et la variance totale. Les trois critères sont équivalents et suivant la situation, l'un peut être choisi. Du fait de la simplicité de la mise en œuvre de la variance interclasse, souvent les auteurs la maximisent pour segmenter des images [11]. Otsu est méthode est un type de seuillage global dans lequel il dépendre uniquement valeur de gris de l'image. Méthode était Otsu proposé par Otsu Scholar en 1979. Otsu méthode est globale Procédé de sélection d'un seuil, qui est largement utilisé parce qu'il est simple et efficace [37]. Le procédé nécessite Otsu le calcul d'un histogramme des niveaux de gris avant de courir. Cependant, en raison de la unidimensionnelle qui ne prendra en considération les informations de niveau de gris, il ne donne pas une meilleure segmentation résultat. Donc, pour que les deux algorithmes Otsu dimensions était proposé, qui fonctionne à la fois sur le seuil de niveau de gris de chaque pixel ainsi que ses informations de

corrélation spatiale à l'intérieur du quartier. Donc l'algorithme Otsu peut obtenir de satisfaisants résultats de segmentation quand il est appliqué aux images bruitées [38].

De nombreuses techniques ont ainsi été proposées pour réduire le temps passé sur calcul et toujours maintenir un seuil raisonnable des résultats. Dans [39], a été proposée une technique récursive qui vise à réduire efficacement le temps de calcul. La méthode d'Otsu était l'une des meilleures méthodes de sélection pour de vraies images du monde à l'égard des mesures d'uniformité et de la forme.

3.5.2 les méthodes paramétriques

Ces méthodes supposent que les niveaux de gris des différentes classes de l'image suivent une certaine fonction de densité de probabilité.

Généralement, ces fonctions de densités de probabilité sont supposées suivre un modèle Gaussien. En partant d'une approximation de l'histogramme de l'image par une combinaison linéaire de Gaussiennes, les seuils optimaux sont localisés à l'intersection de ces dernières.

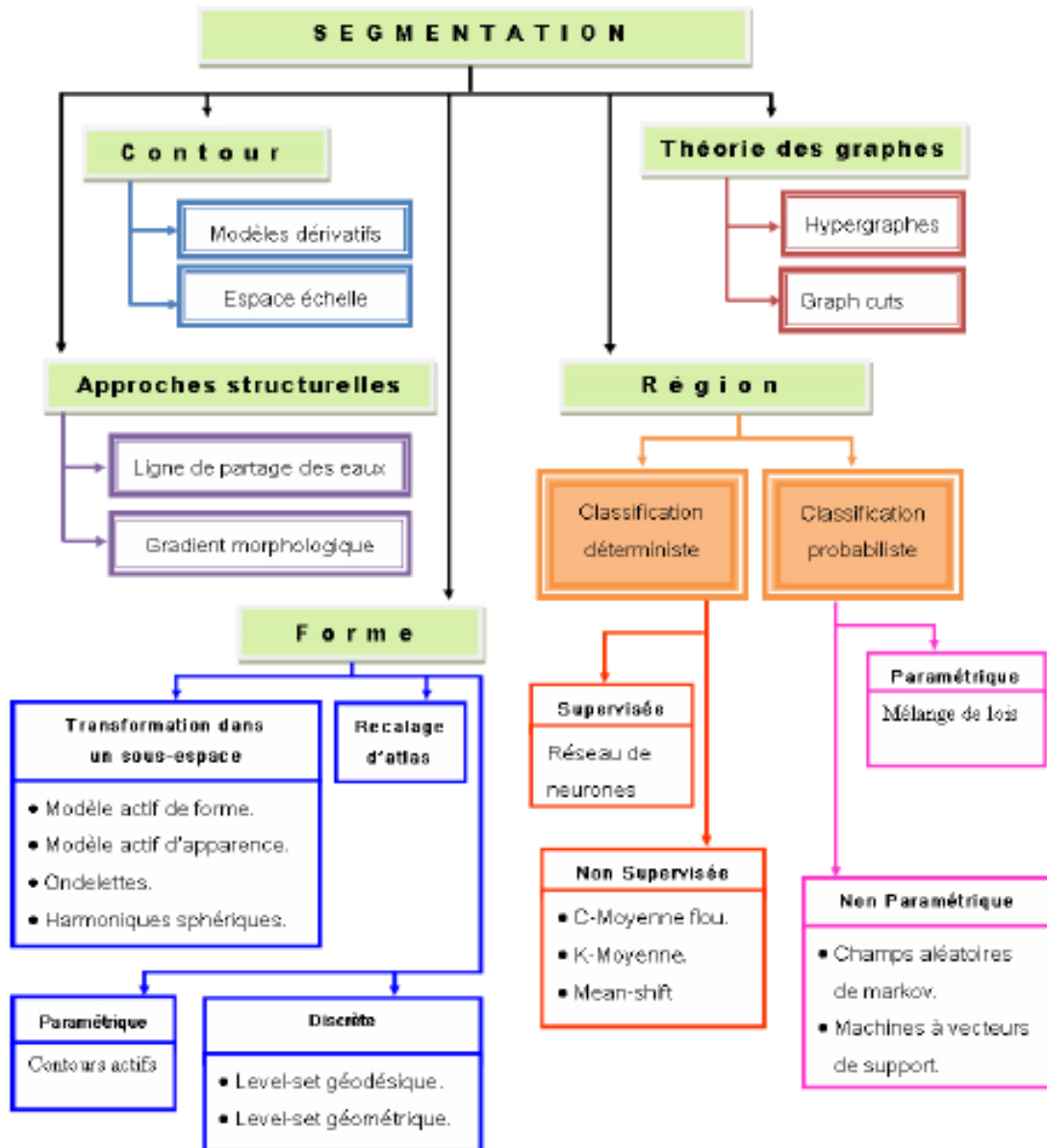


FIGURE 3.4 - Nomenclature des méthodes de segmentation d'image.[30]

3.6 Conclusion

Toutes ces approches visent à l'extraction des indices visuels. Après de nombreuses années passées à rechercher la méthode optimale, les chercheurs ont compris que la segmentation idéale n'existait pas. On peut même montrer que le problème de la seg-

mentation est le plus souvent un problème mal posé. Etant donnée une image, il existe toujours plusieurs segmentations possibles. Une bonne méthode de segmentation sera donc celle qui permettra d'arriver à une bonne interprétation. Elle devra donc avoir simplifié l'image sans pour autant en avoir trop réduit le contenu. Entre autres, elle devra éviter les choix irréversibles. L'avenir de la segmentation est dans le pilotage aval par l'applicatif. [17]

Chapitre 4

Présentation de l'application

4.1 Préliminaire

Dans cette partie expérimentale, nous présenterons les travaux précédents liés à ce thème dans le cadre de l'état de l'art, après nous détaillerons, dans une première contribution le fonctionnement et l'implémentation de la fonction Otsu ainsi que celle de la croissance de région. Et dans une deuxième contribution, nous détaillerons l'emploi de l'indice de circularité, après on verra la méthode élaborée pour limiter la croissance de région. Enfin, Dans la dernière section, nous discuterons les résultats obtenus.

4.2 Etat de l'art

Nous résumons quelques travaux liés à ce thème comme suit :

- Une segmentation de texture par classification pixéllaire qui se base sur les Support Vector Machine (SVM), notre technique permet de faire une sélection d'attributs de texture d'Haralick extrait des matrices de cooccurrence ,sur des images d'apprentissage codées en niveau de gris.
- Une technique qui combine (avec le vote majoritaire et la théorie de Dempster

Shafer) les résultats obtenus par les Support Vector Machine (SVM) dans différents espaces couleur.

- Application en premier lieu d'une analyse discriminant linéaire (*LDA*), pour extraire les composantes de différents espaces couleur (RGB, HSL, HSV, LUV, YUV). En deuxième lieu, l'utilisation de l'algorithme de classification non-supervisée FCM

4.3 Base d'images

Notre base est construite à partir d'images acquises au sein du service d'hémobiologie du C.H.U Tlemcen sur des lames avec la coloration de May Grunwald Giemsa (MGG). L'environnement LEICA (caméra et microscope) permet d'obtenir des images couleur RGB de bonne qualité ; au format BMP de dimension 1024x768.

4.4 L'algorithme d'Otsu

4.4.1 Principe de fonctionnement

La méthode d'Otsu utilise une recherche exhaustive à évaluer le critère de maximisation de la classe, entre variance.[39]

On a choisit la méthode pour les raisons suivantes :

- La méthode d'Otsu est l'une des meilleures méthodes de sélection de seuil pour de vrai images à l'égard des mesures d'uniformité et la forme.
- La méthode d'Otsu utilise une recherche exhaustive à évaluer le critère de maximisation de la classe, entre variance. Comme le nombre dans les classes d'une image augmente,[10].

4.4.2 Implémentation

Il existe plusieurs façons d'implémenter la méthode d'OTSU sur MATLAB. Dans le cas présent, nous allons voir deux méthodes distinctes offrant le même résultat mais avec des caractéristiques différentes. La première méthode consiste à utiliser la fonction « `graythresh` ». Cette méthode est la plus simple car elle nécessite peu de lignes de commande et peu de calcul.

- Chargement de l'image de départ (organigramme)
`Image = imread(Image.jpg);`
- Calcul du seuil T
`level = graythresh(Image);`
- BINARISATION
`BW = im2bw(Image,level);`
- Affichage de l'image

`figure; imshow(BW); title(Image binarisée avec la méthode 1 d'OTSU);` La seconde méthode consiste à implémenter manuellement la méthode d'OTSU. C'est une méthode plus complexe, elle nécessite le calcul de la variance ainsi que de la moyenne de chaque classe pour chaque seuil T possible. Cette méthode est plus calculatoire que la première. On est donc plus amené à faire des erreurs lors de l'implémentation.

Dans notre cas, on a utilisé l'image au canal vert avec une valeur de N qui est égale 7, comme le montre l'instruction suivante :

$$[IDX, sep] = otsu(I(:, :, 2), 7);$$

Après plusieurs essais, on conclut que l'image au canal vert avec un seuillage de sept classes donne le meilleur résultat pour distinguer la classe du noyau qui est la première classe, ceci est illustré dans la figure ci-dessous, sans oublier qu'on peut afficher l'histogramme de l'image au canal vert pour montrer que ce canal contient un pic, qui veut dire que c'est lui qui contient le plus d'intensité et d'énergie.

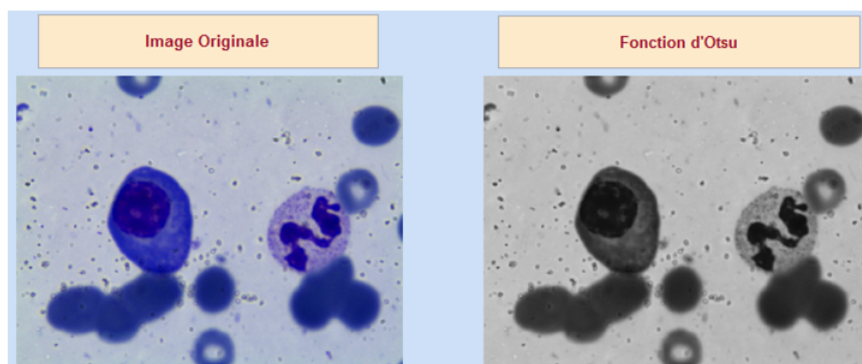


FIGURE 4.1 – Fonction d'Otsu.

[31]

Ci-dessous on va présenter les étapes du programme de la deuxième méthode d'Otsu sous forme d'un organigramme.

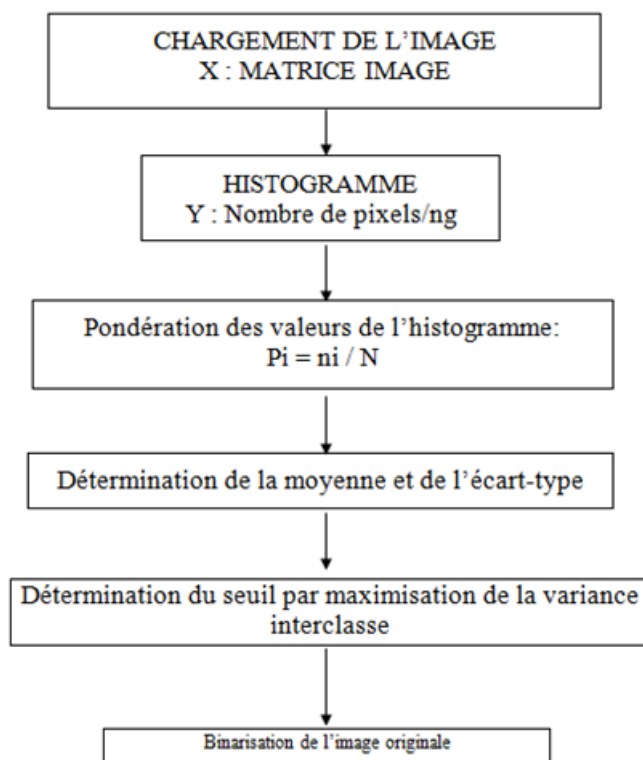


Figure 4.2 - Organigramme de la méthode d'otsu.

* **Comparaison des deux méthodes en terme de temps d'exécution** Le graphique ci-dessous met en évidence le temps d'exécution des deux méthodes en fonction de la taille de l'image.

Pour les deux méthodes, la taille de l'image n'influe pas sur le temps d'exécution. La méthode d'OTSU étant basée sur le calcul de l'histogramme et la taille de l'histogramme ne variant pas en fonction de la taille de l'image, il est donc logique que la taille de l'image n'influe pas sur le temps d'exécution. La méthode utilisant la fonction « graythresh » est moins rapide que l'autre méthode. Cette différence peut s'expliquer simplement par le fait que l'appel de la fonction « graythresh » demande plus de ressources que de parcourir une simple boucle « for » de 1 à 255.

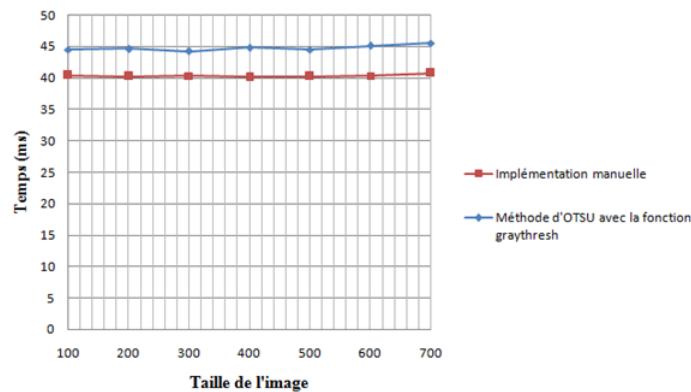


Figure - Temps d'exécution des deux méthodes.

4.5 L'algorithme de Croissance de région

Cet algorithme peut se modéliser sur deux méthodes qui sont les suivantes :

i. Méthode par amorce

Elle s'appelle en anglais « Seed-based region growing » (SBRG) qui a été appliquée profondément dans le domaine médical.

Dans cette méthode on définit une zone R qui contient la région à extraire,

initialement, R contient l'amorce, on utilise une file FIFO (First In, First Out) S qui contient les points frontière, initialement, S contient le voisinage de l'amorce.

On retire p dans S ,

– **Si** p est homogène avec R **Alors**

– On ajoute p à R et on ajoute à S les points du voisinage de p qui ne sont pas dans R et qui ne sont pas incompatibles.

– **Si Non** on marque p comme incompatible.

– **FinSi**

On recommence tant que S n'est pas vide.

ii. Méthode linéaire

Lorsque le critère d'homogénéité est local, alors il y aura une comparaison de la valeur du pixel candidat et du pixel de la frontière, sa complexité est de $2n$ et son algorithme se présente comme suit :

Pour chaque pixel on regarde si sa valeur est égale à la valeur de la région

– **Si** c'est le cas, **Alors**

– Le pixel est intégré à la région.

– **Si Non** Il fait partie de la frontière.

– **FinSi** [40].

4.5.1 Principe de fonctionnement

La technique de croissance de région est une approche locale, et un processus itératif par lequel les régions sont fusionnées à partir de pixels individuels, ou d'une autre segmentation initiale, et la croissance sera de manière itérative jusqu'à ce que chaque pixel soit traité.[29] Supposons une région de couleur homogène R . Initialement,

$$R \geq 1$$

pixel. On va étendre la région R en incluant les pixels situés sur la frontière et dont la couleur est proche de celle de R (la variation de couleur est inférieure à un seuil δ ; caractéristique de ce type d'algorithmes). En répétant cette procédure jusqu'à ce qu'il n'y est plus de pixels de couleur assez proche sur la frontière, on obtient une région de couleur homogène maximale autour du pixel de départ. La région initiale « gonfle » en absorbant des pixels de la frontière, jusqu'à stabilité par rapport à la propriété d'homogénéité i.e Cette méthode est très générale est peut-être appliquée avec n'importe quel critère permettant de décider si l'on doit ajouter ou non un point de la frontière à la région.

1. **Fonctionnement manuel** La plupart des méthodes de croissance de la région identifient les points d'amorce manuellement, ce qui implique l'interaction humaine et exige une information préalable sur l'image, ce fonctionnement contient deux façons :

– **1er cas**

Dans le 1er cas, le principe du fonctionnement manuel est plutôt basé sur l'intégration de deux paramètres x et y représentant les coordonnées d'un pixel préalablement fixé et choisi à partir de la région dont on veut agrandir.

– **2ème cas**

Ce fonctionnement simule le cas où l'utilisateur ne fait pas la saisie de x et y ou bien les deux valeurs seront nulles. Donc ce principe se fait par le biais d'une fonction prédéfinie en Matlab, appelée « `getput(n)` », cette dernière possède un paramètre entier « n » qui doit prendre une valeur de 1 ou plus, et qui précise le nombre de pixels qu'on veut intégrer dans la région initiale. Selon ce nombre de germes, on donne la main à l'utilisateur pour faire autant de cliques sur la région voulue de l'image pour lancer la croissance.

2. **Fonctionnement automatique** Le principe du fonctionnement automatique ne donne pas la main à l'utilisateur de faire la saisie d'aucun des paramètres de la fonction.

Dans ce fonctionnement il y a la méthode par amorce où le point amorce ou germe est choisi de manière automatique en évitant les zones de fort contraste (gradient important).

4.5.2 Implémentation

Il existe plusieurs façons d'implémenter la méthode de croissance de région sur MATLAB. Dans le cas présent, nous allons voir une de ces méthodes qui sont distinctes offrant des résultats différents, sans utiliser les mêmes paramètres, on peut les classer en deux grandes catégories, la première pour les images 2D et l'autre pour les images 3D. Notre méthode fait partie de la première catégorie qui segmente les images de deux dimensions, ci-dessous on va présenter les étapes du programme sous forme d'un organigramme.

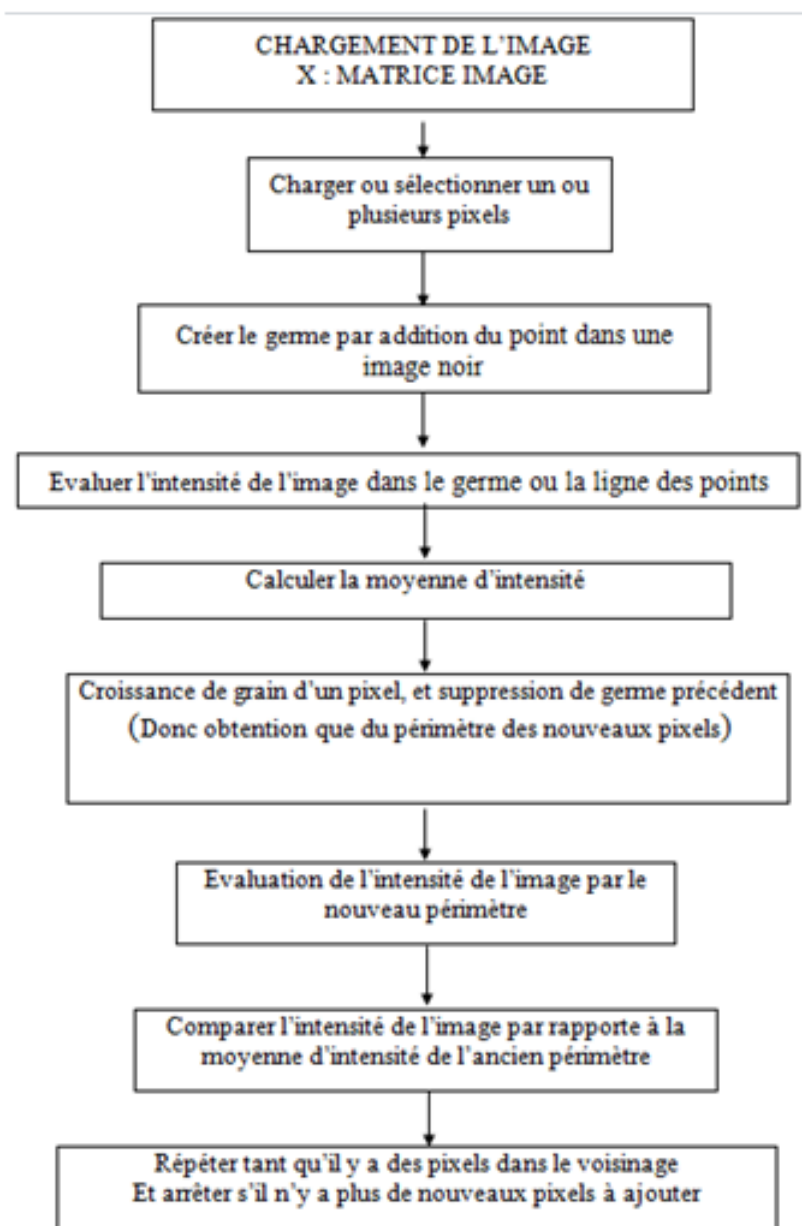


Figure- Organigramme de la croissance de région.

4.5.3 Les paramètres de La fonction

Cette fonction dans le cas classique, possède 3 paramètres qui sont les suivants :

1. L'image I

I est une image couleur RGB pour laquelle on va appliquer la segmentation de

croissance de région.

2. Les germes, amorces, seeds

Les germes sont les points de départ des régions dont on veut segmenter, leur dépôt ne posent pas de problème pour peu que l'utilisateur (physiciens, médecins ...) possèdent un minimum de connaissance anatomique pour pointer la structure qui l'intéresse[41]

3. Le seuil δ

Le critère d'homogénéité a pour effet de favoriser une croissance homogène de la région marquée, en mesurant la différence entre le niveau de gris du pixel et le niveau de gris moyen de la région marquée[42]

Par conséquent, définir l'homogénéité est basé sur des procédures d'essai et d'erreur. Pour cette raison, la recherche basée sur cette méthode s'est concentrée sur la détermination de ces critères aussi bien que sur l'efficacité de l'algorithme. De là, plusieurs approches ont été définies pour obtenir les critères d'homogénéité de façon automatique qu'on classe de trois manières suivantes :

- Les méthodes qui sélectionnent les critères basés sur les niveaux de gris dans le voisinage du germe. Les critères sont mises à jour à chaque fois qu'un point est ajouté à la région. Ces méthodes ont le désavantage d'être dépendent de la position de la germe et de l'ordre de parcours.
- Les méthodes qui sélectionnent les critères basés sur un nombre de différentes segmentations avec des critères d'homogénéité qui varient. Elles sont lentes à cause du nombre de segmentation à effectuer et leur succès dépend de la capacité à distinguer le vrai résultat des segmentations ayant une légère différence de critère.
- Les méthodes qui sélectionnent les critères sur une considération optimale de la

segmentation complète de la scène avec des critères qui peuvent évoluer pour différentes régions[41].

Dans notre cas, nous avons opté pour la troisième méthode qui est une méthode expérimentale, car on a obtenu le seuil après plusieurs essais pour obtenir un seuil fonctionnant pour la plus part des cas.

4.5.4 Avantages et inconvénients de la méthode

i. Avantages

- Elle permet la segmentation d'objet à topologie complexe.
- Elle préserve la forme de chaque région de l'image.
- Conceptuellement très simple
- Méthode rapide
- Elle détache les composantes connexes (dans le cas où deux objets ont une couleur proche)
- La connexité est garantie par la procédure de segmentation, la position et les critères d'homogénéités, donné par l'utilisateur.

ii. Inconvénients

- Une mauvaise sélection des amorces ou un choix du critère de similarité mal adapté peuvent entraîner des phénomènes de sous-segmentation (figure 4.3) ou de sur-segmentation, donc les régions obtenues dépendent fortement des pixels d'amorçage choisis et de l'ordre dans lequel les pixels de la frontière sont examinés.
- Il peut y avoir des pixels qui ne peuvent pas être classés.
- Le résultat final est très sensible à la valeur du seuil δ .

Pour illustrer le premier problème, considérons trois pixels adjacents a, b et c dont les intensités respectives sont 8,10 et 11 (par exemple, l'intensité en niveaux de gris).

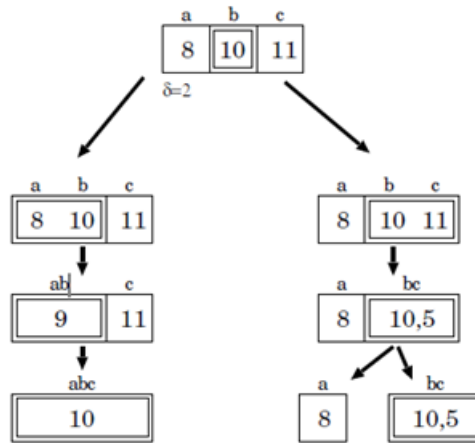


FIGURE 4.2 – Deux schémas de regroupement pour des pixels a, b et c.

[31]

Le seuil est 2, la région initiale est constituée du pixel b. Deux schémas de regroupement pour les points frontière a et c sont possibles (figure 4.1)

Le pixel central b est l'amorce. Compte tenu du seuil $\delta = 2$, a et c sur la frontière devraient être ajoutés à l'amorce b. Si l'on commence par tenter d'agréger a, le résultat du regroupement, noté [ab], a pour intensité moyenne 9 et c s'y ajoute ensuite puisque On a donc regroupé a, b et c. Si maintenant l'algorithme commence par examiner le point frontière c, le groupement de b et c donne [bc] dont l'intensité est 10; 5. Le point frontière a d'intensité 8 est trop éloigné et il est considéré comme appartenant à une autre région. On obtient donc deux groupements au lieu d'un.

Ce petit exemple illustre combien l'ordre d'examen des points sur la frontière peut influencer sur le résultat de l'algorithme. A fortiori, cet algorithme est également très sensible au choix des amorces.

Par ailleurs, comme nous allons le voir, le résultat final dépend très sensiblement du seuil. Une petite variation de ce seuil peut conduire à des modifications importantes.

Cet algorithme fait partie de la classe d'algorithmes de segmentation dit « locaux

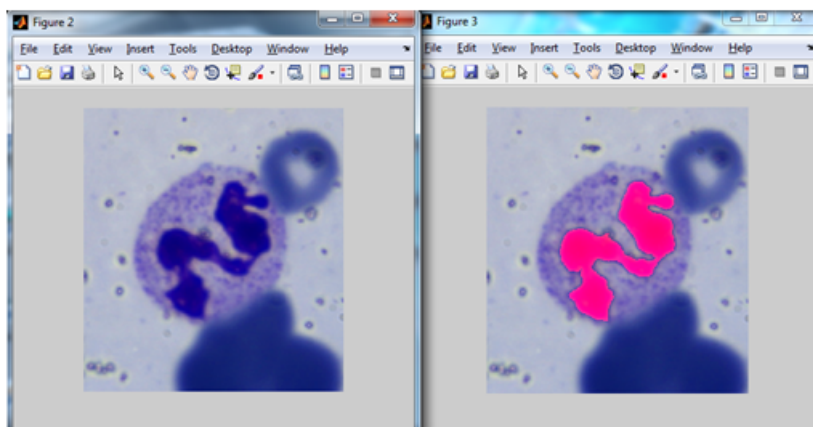


FIGURE 4.3 – Exemple de sous segmentation avec un seul point d'amorce.

». L'opération élémentaire consiste à manipuler des pixels adjacents et l'algorithme n'a aucune vision globale du résultat qu'il obtient. Par exemple, il est incapable de détecter que la région qu'il vient de construire est inhomogène.[26]

4.5.5 Solutions proposées

Pour faire face aux limites qu'on peut trouver dans cette fonction, on a réalisé quelques contributions qu'on va citer ci-dessous :

- Extraction des pixels de la région du noyau pour lancer la croissance
- Recherche d'une valeur optimale du seuil pour donner les meilleurs résultats
- L'ajout de la boucle pour améliorer l'affichage, simplifier l'interprétation et voir dans quel sens se fait la croissance.
- Intégration de l'indice de circularité.
- Test d'arrêt pour limiter la croissance.

Par la suite, on va détailler ces grands points, après avoir expliqué quelques démarches et traitements importants sur les images qui sont inclus dans notre interface.

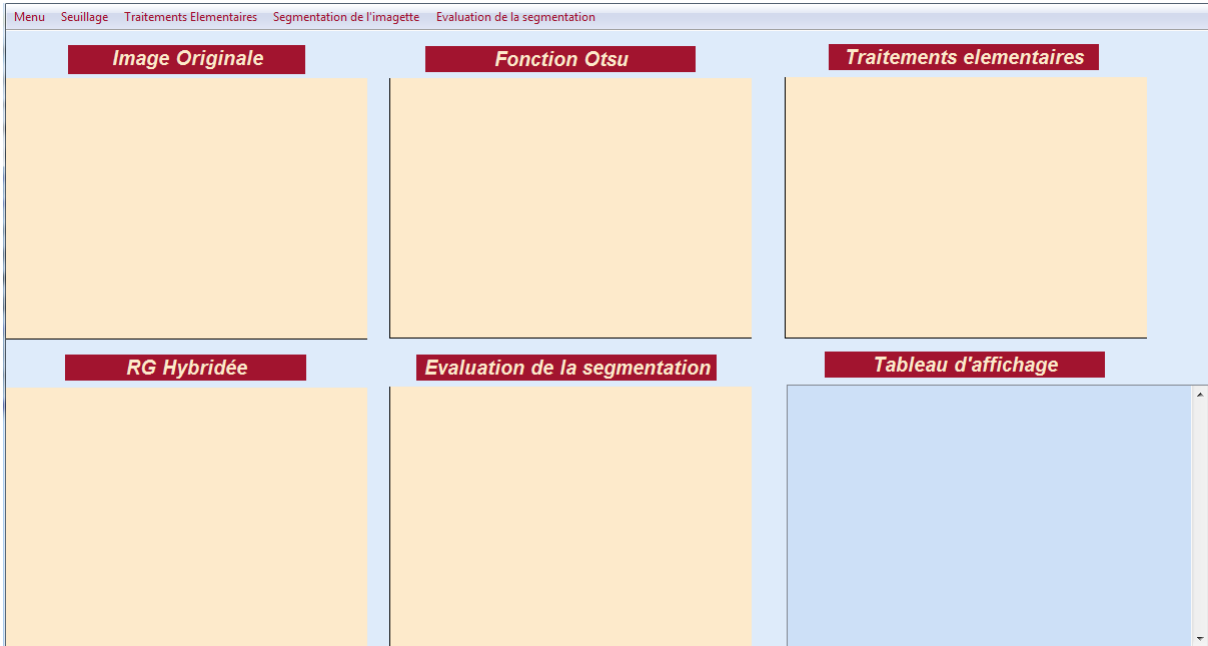


FIGURE 4.4 – Interface de l'application.

4.6 Interface de l'application

Nous exposerons l'interface de notre application, en essayant à chaque fois de décrire les différentes fonctionnalités mise à la disposition de l'utilisateur, et d'expliquer les étapes qui ont précédées l'obtention du résultat final.

Lors du lancement de l'application (figure 4.4) il faut tout d'abord commencer par faire le chargement de l'image, et cela se réalise à l'aide de l'option "menu" de notre interface.

Après avoir chargé l'image, on lui applique la fonction Otsu (figure 4.5) avec les paramètres nécessaires, ensuite on fait une labélisation de la première classe qui simule les noyau en leur affectant l'étiquette 1 qui indique la couleur blanche dans une image binaire, et on laisse le reste des objets de l'image en noir, cela est illustré dans la figure 6.2. L'étape qui suit l'extraction binarisée de tous les noyaux existant dans l'image, et l'élimination du bruit, qui consiste à se débarrasser de tous les éléments non noyaux et cette phase nécessite une condition que nous avons proposés pour donner des résultats

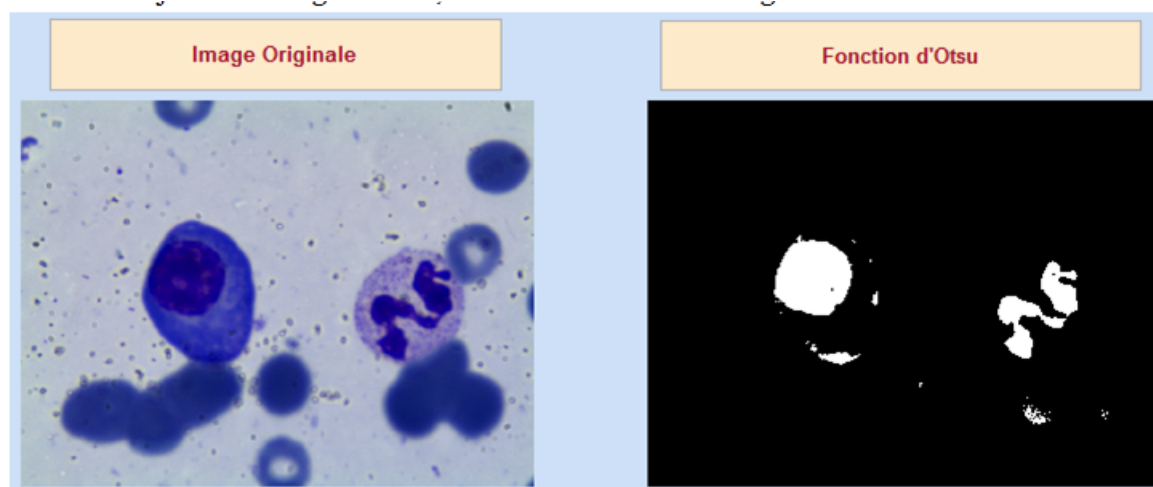


FIGURE 4.5 – Extraction originale des noyaux.

plus fiables, la condition est la suivante :

Si le nombre d'élément avec un label=1 est supérieur à 2

Alors supprimer toutes les régions inférieures à un certain nombre de pixels (estimé expérimentalement à 5500 pixels)

Sinon supprimer toutes les régions moins de 2000 pixels (La suppression se fait à l'aide de la fonction de Matlab « `bwareaopen` »).

Dans l'étape qui suit, on fait le remplissage dans le but de remplir le vide dû à des pixels mal classés au sein du noyau, ainsi que pour joindre les régions du noyau faiblement liées qui est le cas des cellules PNN et PNE, pour ne pas avoir des cas de conflits dans le comptage (le comptage se fait avec la fonction « `bwlabeln` » de MATLAB) à cause de la perte de quelques pixels reliant ces régions.

La phase de dilatation consiste à élargir le périmètre du noyau, et le rendre purement convexe dans les zones concaves, et cela va augmenter sa superficie pour gagner en terme de temps d'exécution dans la méthode de croissance de région, donc la région initiale contiendra à l'avance une partie du cytoplasme.

L'étape qui suit la dilatation est l'affichage de chacun des noyaux existants dans la

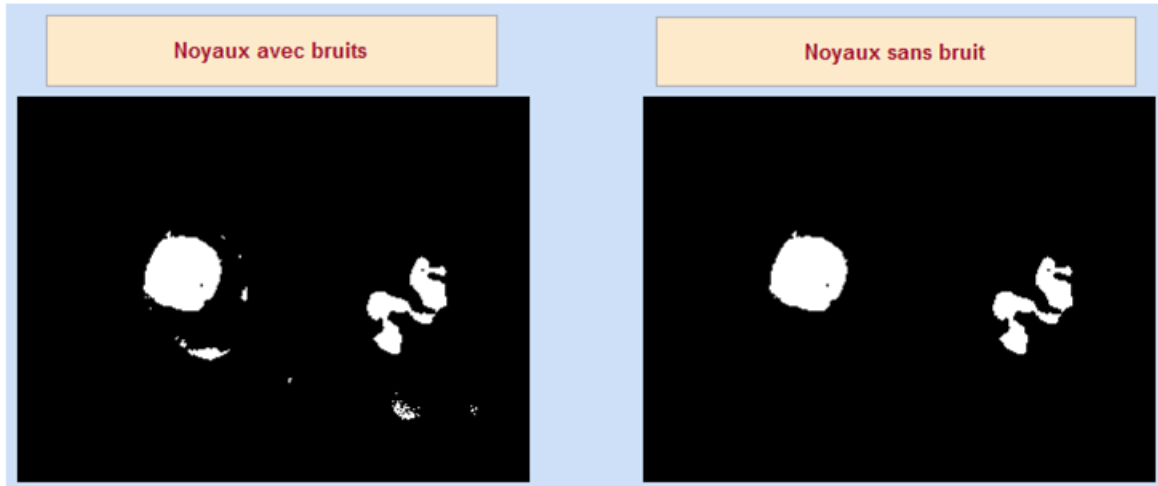


FIGURE 4.6 – Elimination du bruit.

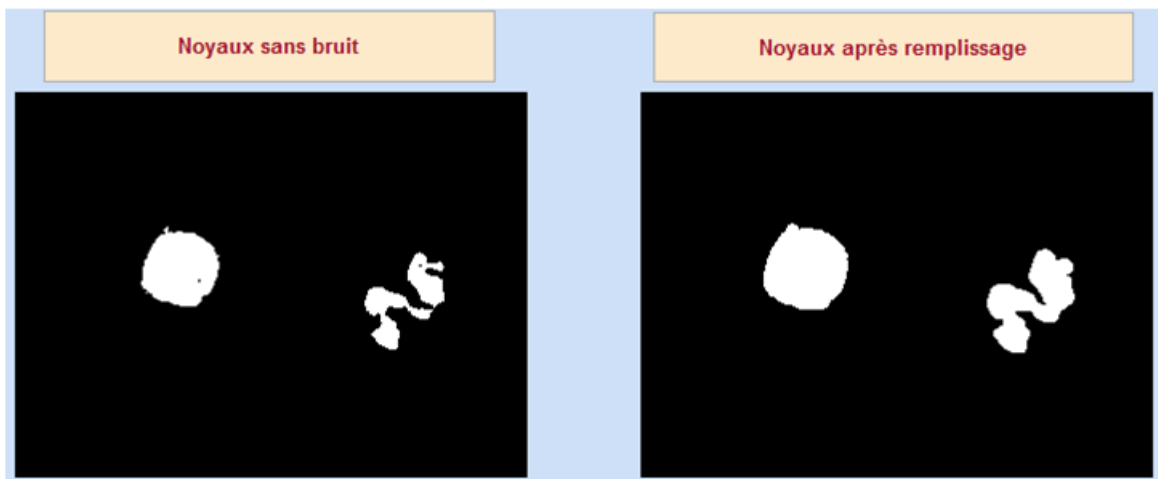


FIGURE 4.7 – Phase de remplissage.



FIGURE 4.8 – Noyaux convexes et dilatés.

capture du frottis dans une image séparée appelée imagette, cette opération est faite à partir d'un comptage du nombre d'objets étiquetés à 1 dans l'image binaire, le chiffre résultant de ce comptage va limiter le nombre d'itérations dans la boucle qui va ajouter une centaine de pixels dans chacune des quatre extrémités du noyau, et cela dans le but d'avoir une imagette contenant l'intégralité du globule blanc.

Cela est résolu à l'aide des fonctions min et max qu'on a appliqué pour chaque pixel du noyau, bien sûr l'ajout des pixels se fait après chaque vérification de ne pas avoir dépassé la taille de l'image.

Sans oublier que les principales fonctions de Matlab qu'on a utilisé pour réaliser tout ce traitement sont : `imclose`, `strel`, `imdilate` et `bconvhull`.

Par la suite, on fait un enregistrement automatique pour chacune des imagettes de globules obtenues, car tout le traitement précédent qu'on vient de citer va se refaire pour les nouvelles images.

La dernière phase qui suit, on applique l'algorithme de croissance de région, et l'appel de la fonction sera comme suit :

```
Phi = segCroissRegion(seuil,image,image binaire du noyau);
```

Les premières modifications qu'on a appliqués sur cette fonction concernent :

- L'automatisation de cette dernière en éliminant les paramètres x et y , et en les remplaçant par le nouveau paramètre « nbw » qui est le résultat du seuillage de la méthode Otsu.
- L'augmentation du seuil de variance jusqu'à 50 pour permettre l'intégration de quelques pixels dans la classe du cytoplasme, qui étaient mal classés à cause de la différence de couleur dû à la coloration du frottis, et à la texture du cytoplasme.
- L'ajout d'une nouvelle boucle « for » au sein de la boucle « tant que » du programme de croissance de région, c'est cette dernière qui s'occupe de la comparaison entre les nouveaux pixels et les pixels déjà classés qui appartiennent au cytoplasme. Cette boucle se lance dès les premières itérations de la boucle « tant que », et à chaque itération on affiche le résultat des pixels récemment classés, ce qui nous permet de voir comment la croissance s'étale sur la région du cytoplasme.

4.6.1 Discussion du premier résultat

Notons que le résultat de cette fonction était de couleur binaire, alors on a appliqué la fonction "label2rgb" pour obtenir une couleur, et cela pour une meilleure interprétation et plus de signification.

Dans le cas où le globule blanc est séparé de toute autre cellule, on remarque que la fonction de croissance de région donne de meilleurs résultats, car on arrive à très bien distinguer le cytoplasme de la classe inter cellulaire et facilement tracer le contour du globule (figure 4.10).

Dans le deuxième cas, les résultats obtenus sont satisfaisants, nous remarquons néanmoins qu'une région appartenant à une cellule voisine a été détectée comme cytoplasme à cause de la similitude dans leurs couleurs, ce qui est le cas d'une sur-segmentation.

Pour une première solution, nous avons essayé de chercher une valeur minimale du seuil qui puisse séparer ces deux classes. Mais, à notre avis, cette solution n'a pas

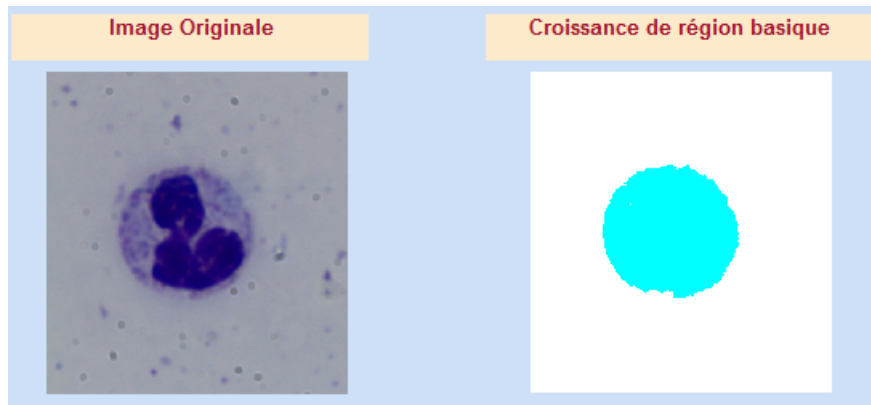


FIGURE 4.9 – Le résultat de la croissance de région (1er cas).

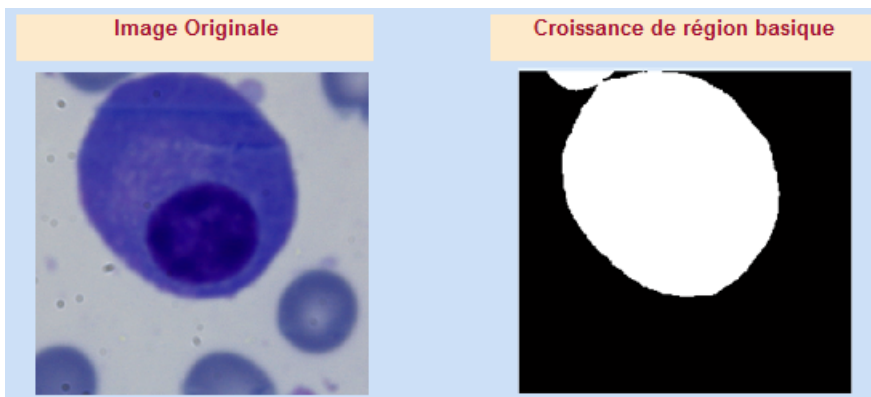


FIGURE 4.10 – Le résultat de la croissance de région (2ème cas).

donné de bons résultats puisqu'une grande réduction du seuil ne permet même pas à la croissance de toucher le cytoplasme.

4.6.2 La phase d'hybridation

Afin de mieux cerner notre sujet, nous avons jugé nécessaire d'approfondir nos connaissances dans le domaine de cytologie. Ainsi les informations que nous avons pu récolter durant le stage telle que de petites recherches théoriques effectuées sur l'Atlas médical dans le domaine de la cytologie ou quelques manipulations par le microscope, sur des échantillons de frottis et clôturé par de longues discussions avec des médecins spécialistes dans ce domaine. Nous avons pu apporter la confirmation que les leucocytes cités dans le premier chapitre possèdent une forme géométrique ronde ou elliptique.

Cette information nous a conduits à chercher une solution pour le problème suivant : comment intégrer la circularité des globules blancs dans le domaine de la segmentation ? A l'aide d'une multitude de recherche dans les calculs mathématiques des cercles, nous avons trouvé que le mathématicien Miller (1953) avait bien défini son indice de circularité, appelé aussi indice d'allongement ou de compacité, et dont la conception caractérise la circularité de la forme étudiée. Il est à noter que l'indice de circularité de Miller est donné par la formule suivante : « le rapport de la surface, au carré du périmètre multiplié par la constante 4π »

La réalisation de l'hybridation entre l'indice de circularité et la fonction de croissance de région nous a permit de faire face à la sur-segmentation et de limiter la croissance de région.

Cette deuxième contribution nécessite quelques calculs avant de l'entamer, comme celui du périmètre et de la surface. Le périmètre a été calculé comme étant une propriété de la région courante à l'aide de la fonction « *regionprops* », cette dernière nécessite donc deux paramètres qui sont l'image et le libellé « *perimeter* » qui indique le périmètre, mais le calcul de l'aire de l'objet a été fait par la fonction « *bwarea* ». Ensuite nous avons créé une variable nommé « indice » qui calcule le rapport et nous affecte le résultat final.

Le paramètre de rondeur est très rapide d'acquisition. Cet avantage nous a permis de le calculer à chaque itération de la boucle « *Tant que* », mais cela n'aura pas une grande influence sur le temps d'exécution.

Le but d'utiliser l'indice de circularité est d'inspecter la façon dont la région s'élargie pour savoir exactement quand est ce que nous devons arrêter la croissance. Nous nous basons sur le principe que l'indice de circularité du globule blanc, car il commence à diminuer quand il y a une intégration des pixels de la cellule voisine, donc nous pouvons conclure que cette cellule a commencée à perdre son architecture ronde.

Le résultat de cette opération est compris entre 0 et 1. Vaut 1 si l'objet est un cercle, 0 si il est de surface nulle. La liste complète des indices de circularité est en annexe de ce document.

En se basant sur ce principe, nous avons créé un tableau qui prend les résultats de l'indice de circularité à chaque itération. Ensuite nous comparons chaque valeur avec celle qui la précède, si cette dernière est supérieure c'est qu'il y a une déformation dans sa forme, alors nous l'arrêt de la croissance à l'aide de la fonction « *Break* » qu'on va inclure dans la boucle « *Tant que* ».

Puisqu'on lance la croissance de région à partir d'une première segmentation à l'aide de la fonction d'Otsu, et après avoir rendu ces noyaux résultants convexes et dilatés (figure 4.8). Cela veut dire que leur forme est proche du cercle donc pour un premier calcul de l'indice de circularité, sa valeur sera élevée, et cette déduction nous a poussé à laisser la croissance s'élargir, ce qui veut dire qu'on doit attendre quelques itérations avant de commencer le teste de comparaison.

Dans la figure ci-dessus, l'indice de circularité à atteint une valeur égale à 0.9451 et dans l'ittération qui suit, il a diminué à une valeur égale à 0.9425, et cette valeur a été la dernière valeur calculée par notre fonction avant de l'afficher, cela prouve que le teste de comparaison s'effectue à chaque ittération.

Ce résultat est considéré comme l'un des meilleurs résultats de la fonction après avoir intégré l'indice de forme.

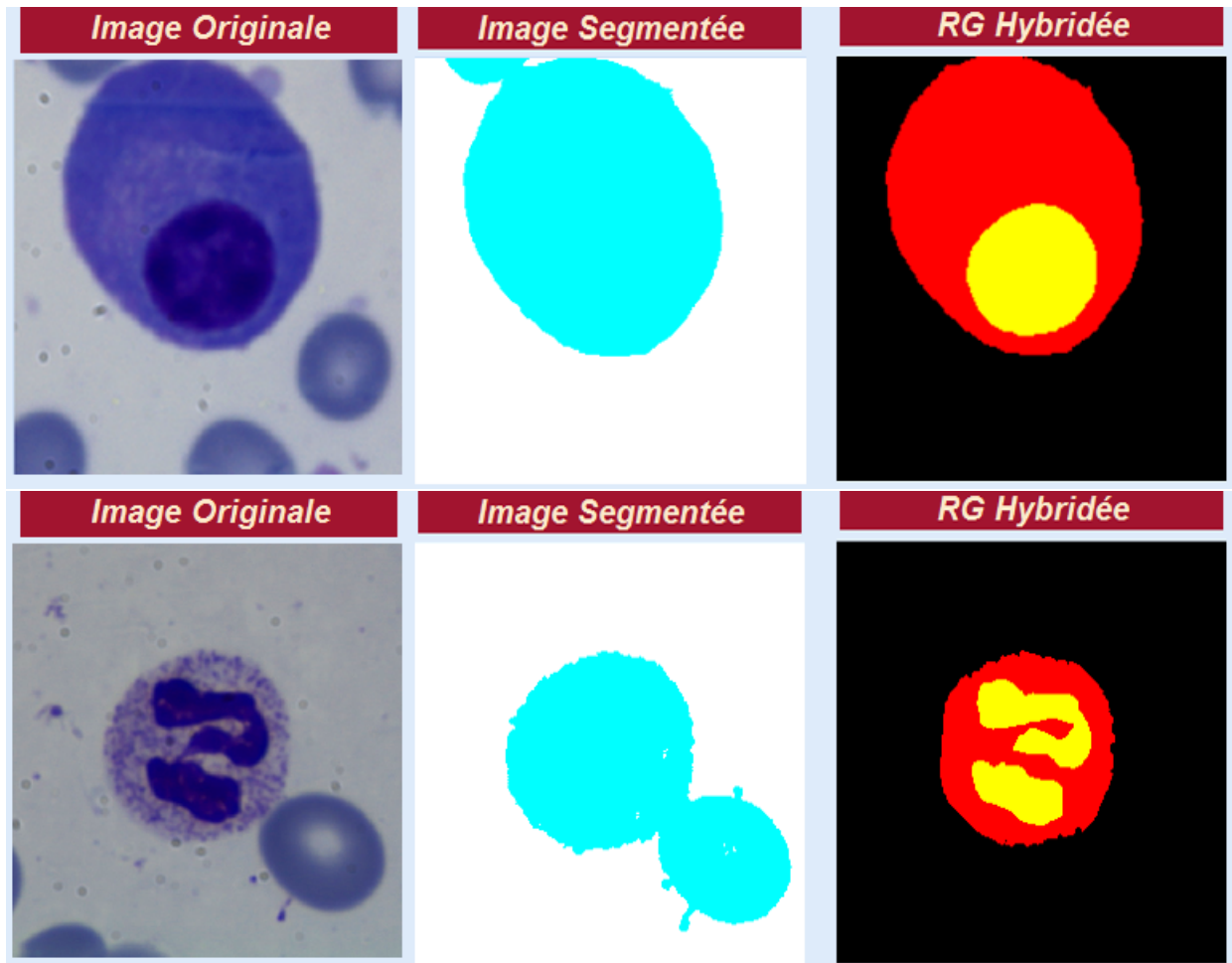


FIGURE 4.11 – Résultat de la fonction hybridée.

4.6.3 Evaluation de la segmentation

Les critères d'évaluation quantitative découlent de deux groupes, à condition de disposer ou non d'une "vérité-terrain" qui constitue une segmentation de référence.

L'évaluation visuelle de résultats de segmentation reste toujours présente et particulièrement dans domaine médical, mais présente l'inconvénient lié à sa mise en œuvre qui s'avère particulièrement longue et fastidieuse pour les experts. Il devient alors primordial de procéder à une évaluation complètement automatisée.

Nous allons présenter une petite définition pour les deux groupes précédemment cités, ensuite nous proposons une nouvelle méthode d'évaluation spécialement adaptée aux attentes des experts du domaine de la cytologie médullaire.

A. Sans vérité-terrain

Ces évaluations cherchent à quantifier la qualité ou la lisibilité de l'image, ils ont l'avantage de ne pas nécessiter de vérité terrain ou de segmentation de référence. Leur utilisation est justifiée dans des domaines d'application où la construction d'une base de vérité terrain est délicate ou trop coûteuse en temps.

B. Avec vérité-terrain

Acquérir une segmentation de référence ou d'une vérité terrain est parfois difficile pour plusieurs causes. La première réside dans l'incapacité d'obtenir un nombre suffisant d'images pour constituer une base de référence du fait dans certains cas de la nature de ces images, la deuxième est liée au problème de concordance entre les experts (absence d'une harmonie), ou encore il est souvent pénible de délimiter les pourtours de régions (présence de flou lors de l'acquisition, de bruit, de débordement de coloration) et pour finir le temps excessif que doit passer un expert pour segmenter finement une seule image. Étant donné qu'on dispose d'une segmentation de référence, les méthodes d'évaluations seront basées sur une mesure de similarité entre l'image segmentée et sa vérité terrain.



FIGURE 4.12 – Comparaison et similarité entre le résultat et l'étiquette.

C. Une méthode d'évaluation adaptée à la cytologie médullaire Dans un système d'aide à la décision et particulièrement le domaine médical, la méthode d'évaluation des résultats est capitale pour établir un diagnostic fiable. À cet effet, nous allons établir une méthode d'évaluation spécialement adaptée à la cytologie médullaire. Puisque notre segmentation des images cytologiques médullaires s'inspire de l'approche de segmentation par classification pixellaire, il convient alors d'opter pour une évaluation de segmentation basée sur des évaluations des performances du classifieur.

Notre travail est au fait une continuité aux travaux de recherche réalisés précédemment, dans lesquels la classification pixellaire a été adoptée. Nous avons réalisé cette dernière basée sur un classifieur supervisé. On rappelle aussi que nous disposons d'une base (images, étiquettes) qui nous permet d'accomplir notre classification de manière convenable. Par conséquent, nous présentons ci-dessous les critères d'évaluation les plus communément utilisés.

On se base sur un appariement des pixels des images segmentées et vérité terrain, et on calcule le nombre de pixels correctement identifiés et les pixels mal classés. On entend par un mal classé, soit un pixel de la classe 3 dans l'image de référence n'est pas reconnu comme étant de la même classe dans l'image segmentée, ou bien un pixel n'étant pas de la classe 3 dans l'image segmentée est reconnu comme étant de la classe 3

dans l'image de référence. On définit dans un test de classification binaire, un

résultat est dit :

- Vrai Positif (TP : True Positif) lorsqu'un item est correctement détecté par le test ;
- Faux Positif (FP : False Positif) est le résultat pour un item déclaré positif alors qu'il ne l'était pas ;
- Faux Négatif (FN : False Negatif) pour un item déclaré négatif alors qu'il était en réalité positif ;
- Vrai Négatif (TN : True Negatif) est l'item correctement déclaré comme négatif.

D. Matrice de confusion deux classes

		Réel	
		Positive	Negative
Prédit	Positive	TP	FP
	Negative	FN	TN

[5]

Pour comparer les résultats avec les images vérité terrain, il faut tout d'abord étiqueter chacune des ses trois classe, à l'aide de la fonction "*classperf*", car si on injecte directement les images de la figure précédente, on n'aura une mauvaise comparaison, car quelques soit le nombre de classes existants dans l'image la matrice de confusion sera de dimension 2.

qui sont numérotées comme suit

1. la première classe désigne le noyau du globule blanc, et c'est celle qui comporte généralement le minimum de pixel.
2. la classe du cytoplasme possède le numéro deux.
3. Dans la troisième classe nous avons mit toute les cellules restantes y compris les globules rouges, donc c'est la classe qui comporte le plus de pixel.

A partir du résultat que nous venons de présenter, on peut tirer les matrices de confusion et toutes les valeurs nécessaires à l'analyse des résultats. On remarque que le taux de reconnaissance à atteint les 96%, donc ce pourcentage représente la valeur absolue des pixels bien classés, alors que le taux d'erreur est vraiment négligeable du fait qu'il ne dépasse pas les 0.4 %, donc on peut

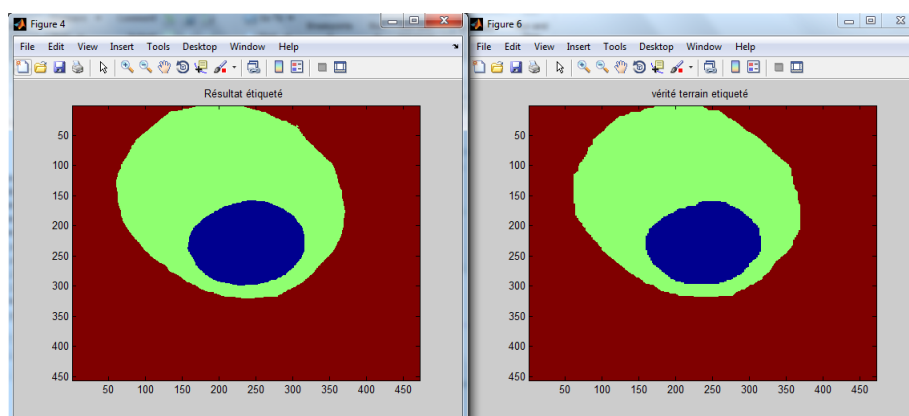


FIGURE 4.13 – Classification et comparaison.

```

Label: ''
Description: ''
ClassLabels: [2x1 uint8]
GroundTruth: [645696x1 double]
NumberOfObservations: 645696
ControlClasses: 2
TargetClasses: 1
ValidationCounter: 1
SampleDistribution: [645696x1 double]
ErrorDistribution: [645696x1 double]
SampleDistributionByClass: [2x1 double]
ErrorDistributionByClass: [2x1 double]
CountingMatrix: [3x2 double]
CorrectRate: 0.9628
ErrorRate: 0.0372
LastCorrectRate: 0.9628
LastErrorRate: 0.0372
InconclusiveRate: 0.2444
ClassifiedRate: 0.7556
Sensitivity: 0.8503
Specificity: 0
PositivePredictiveValue: 0.8343
NegativePredictiveValue: 0
PositiveLikelihood: 0.8503
NegativeLikelihood: NaN
Prevalence: 0.8555
DiagnosticTable: [2x2 double]
    
```

FIGURE 4.14 – Les résultats de Classperf.

conclure que notre objectif a été atteint.

4.7 Conclusion

Dans ce chapitre, nous avons essayé de justifier le choix de méthodes dont on a travaillé avec. Ainsi que nous avons présenter leurs points forts et aussi nous avons traité des petites lacunes de la fonction de croissance de région. Pour conclure, sur la totalité des images testées et sur toutes les segmentations faites, nous avons constaté que dans la majorité de images on arrive à stopper la segmentation qui dépasse les limites du cytoplasme, malgré l'existence de quelques cas difficiles à satisfaire.

Conclusion générale et perspectives

L'hybridation entre une méthode de segmentation dans l'analyse intelligente d'images et les indices de formes basés sur des calculs mathématiques est possible. Elle est réalisable dans le domaine du traitement des images microscopiques issues de frottis sanguins. Le traitement d'image que nous avons pu mener grâce à la disponibilité d'une base contenant une trentaine d'images par grossissement 100 à l'aide du microscope optique très performant du type (*Leica*), nous a permis :

1. La détection de tous les noyaux de leucocytes présents dans chacune des images traitées ;
2. L'élimination du bruit pouvant être confondu avec le noyau ;
3. Le comptage des globules blancs présents dans l'image.

L'intégration de l'indice de circularité dans la méthode de croissance de région à montrée que :

1. L'obtention de la région initiale à accroître peut être réalisé à l'aide d'un seuillage avec la fonction Otsu ;
2. Le Calcul itératif de la circularité d'un leucocyte présent pendant la croissance de sa région a été effectué ;
3. Il est possible de contourner le gros inconvénient de la sur-segmentation.
4. Le temps d'exécution de notre programme est de l'ordre d'une dizaine de secondes comparativement à d'autres segmentations qui dureraient quelques minutes.

Dans la continuité directe de notre travail, nous avons jugé intéressant de proposer ces divers points de recherche pouvant faire l'objet d'étude à l'avenir :

- L'utilisation de la méthode linéaire dans la croissance de région ;
- L'utilisation de la fonction "grayThresh" dans le seuillage d'Otsu ;
- l'élaboration d'un bilan à travers la caractérisation des globules blancs dans le but de faire le comptage des différents types.

Bibliographie

- [1] A.Quessar. Cours d'hématologie. 2015.
- [2] Hémogramme : Indications et interprétations. 2010.
- [3] H.Jouault N.Lefevre-Bultingaire. Frottis sanguin. *Université de Montpellier*, 2010.
- [4] Dr Ouslim. Frottis sanguin. 2010.
- [5] A. Charpentier. Myélogramme normal chez l'adulte. 2013.
- [6] F. Valensi. Morphologie des cellules sanguines normales. 2006.
- [7] technoprimaire. [http :/www.adresse.com](http://www.adresse.com), consulté le 12/03/15.
- [8] [http :/www.maverick.inria.fr](http://www.maverick.inria.fr), consulté le 30/04/15.
- [9] O.Monga. Segmentation d'images par croissance hiérarchique de régions. 1988.
- [10] Y.GAGOU. Cours de traitement d'image. 2008.
- [11] N.Rabia. revue des méthodes de segmentation d'images texturées : cas des images couleur. *Magister en électronique*, juin 2011.
- [12] L.BUSIN. sélection automatique d'espace couleur pour la segmentation d'images application à la qualité de motifs de verres sérigraphies. *thèse de doctorat*, 2006.
- [13] S.OUATARA. stratégies de segmentation d'images multi-composantes par analyse d'histogrammes multidimensionnels. *thèse de doctorat*, 2009.
- [14] N. Vandenbroucke. segmentation d'images couleur par classification de pixels dans des espaces d'attributs colorimétriques adaptés. application à l'analyse d'image de football. *thèse de doctorat*, 2000.
- [15] Y.Gousseau et al I.bloch. le traitemet des images. 2005.

- [16] S. W. Zucker. Region growing : childhood and adolescence. 1976.
- [17] Segmentation-images-principes. [http :/www.glotin.univ-tln.fr/MCBIR](http://www.glotin.univ-tln.fr/MCBIR), consulté le 29/04/15.
- [18] A.Porebski. Classification d'images couleur texturées. *Rapport de master*, juin 2006. [http ://glotin.univ-tln.fr/MCBIR](http://glotin.univ-tln.fr/MCBIR), consulté le 29/04/15.
- [19] I.Baghli. Fusion de segmentation d'images cytologiques par svm dans différents espaces couleur. *mémoire de master*, 07 Juillet 2011.
- [20] S.L. Horowitz and T. Pavlidis. Picture segmentation by a directed split and merge procedure. pages 101–11, 1977.
- [21] Priese and V. Rehrmann. A fast hybrid color segmentation method. in s. j. poppl and h. handels. pages 27–29, 1993.
- [22] M. GARCIA L.SANT'ANNA, L.BINS. Satellite imagery segmentation : a region growing approach.
- [23] M. Y. Mashor² N. A. Mat-Isa¹ and N. H. Othman³. Seeded region growing features extraction algorithm.
- [24] S.L. Horowitz and T. Pavlidis. Picture segmentation by a directed split and merge procedure. pages 101–11, 1977.
- [25] L. Li J. Gong and W. Chen. Fast recursive algorithms for two-dimensional thresholding. pages 295–300, 1998. Pattern Recognition.
- [26] L. G. Roberts. Machine perception of three dimensional solids. pages 159–197, 1965. Optical and Electro-optical Information Processing.
- [27] J. M. S. Prewitt. Object enhancement and extraction. pages 75–149, 1970.
- [28] I. Sobel. Neighbourhood coding of binary images for fast contour following and general array binary processing. pages 127–135, 1978. Computer Graphics and Image Processing.
- [29] K.LARBI. Segmentation d'images basée sur la modélisation statistique d'histogrammes. *Mémoire de majister*, 2012.

-
- [30] M. Benazzouz. analyse intelligente des images médicales : application aux images microscopique de cytologie. *Thèse de Doctorat*, Janvier 2014.
- [31] N. Otsu. A threshold selection method for grey level histograms. 1979.
- [32] T. Pun. A new method for gray-level picture thresholding using the entropy of histogram. pages 223–237, 1980.
- [33] P. K. Sahoo E. J. N. Kapur and A. K. C. Wong. A new method for gray-level picture thresholding using the entropy of histogram. pages 273–285, 1985.
- [34] G. Johannsen and J. Bille. A threshold selection method using information measures. *Conf. On Pattern Recognition*, 1983.
- [35] C. H. Li and C. K. Lee. Minimum cross entropy thresholding. *Pattern Recognition*, pages 617–625, 1993.
- [36] C. Wilkins P. Sahoo and J. Yeager. Threshold selection using renyi’s entropy. *Pattern Recognition*, pages 71–84, 1997.
- [37] Zhong Qu and Li Hang. Research on image segmentation based on the improved otsu algorithm. *Pattern Recognition*, 2010.
- [38] Li Wen-qing LIU Jian-zhuang. The automatic threshold of gray level pictures via two-dimensional otsu method. *Acta Automatic Sinica*, 1993.
- [39] L. Li J. Gong and W. Chen. Fast recursive algorithms for two-dimensional thresholding. *Pattern Recognition*, page 295–300, 1998.
- [40] manzaner. Jcb-segmentation. *Cours*. <http://perso.ensta-paristech.fr/~manzaner/Cours/D9-2/JCB-Segmentation.pdf>, consulté le 19/03/15.
- [41] Klaus D. Toennies R. Pohle. Self learning model-based segmentation of medical images.
- [42] L. Djerou. une approche biomimetique pour la segmentation d’images. Mars 2009.