

17/004.410

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université 08 Mai 1945-Guelma

Faculté des Mathématiques, d'Informatique et des Sciences de la matière

Département d'Informatique



Mémoire de fin d'étude de Master

Filière : Informatique

Spécialité : Ingénierie des Medias

12/798

Thème :

Interprétation automatique des anomalies dans les images médicales

Sous la direction de :

Mr. Ferkous Chokri

Présenté par :

Soltani Charif,

Bendjama Tarek

Juin 2012



Remerciements :



On tient à remercier en premier lieu mon **Dieu** le clément qui nous a guidé dans mon parcours.

On tien à exprimer notre profonde gratitude et nos sincères remerciements pour notre encadreur :

« Mr : Ferkouse Chokri »

Pour son aide précieuse et ses conseils au cours de l'élaboration de cc travail.

On adresse mes vifs remerciements aux membres du jury pour m'avoir fait l'honneur d'évaluer mon travail.

Nos remerciements vont également à l'ensemble des enseignants du département d'informatique pour leurs aides durant toutes mes années d'études.



Dédicace :

Je dédie ce modeste travail à mes très chers parents : « Moussa » et « Samia » qui ont sacrifiés leur vie pour donner sens à la mienne.

A mes chères sœurs :

« Soumeya, Imen»

A ma chère amour et femme Hadjer qu'elle ma donné le courage ,

A mes amis : « Salah, Ali, Assaad, Zaki, Fares, Sami»

A mes collègues du travaille « Hako, Raouf, Edaragé, samir , salah , , , , , a tous les travailleurs du contrôleur financier »

A tous ceux qui ont partagés avec moi mes années d'études et les meilleurs moments de ma vie.

Tarek

Dédicace :

الحمد لله على كل شيء أولا

Je dédie ce modeste travail à mes très chers parents : « Abide » et « Rachida » qui ont sacrifiés leur vie pour donner sens à la mienne.

Mes oncles et mes tantes et mes grands pères et mes grandes mères

A ma chère sœur « khawla » et ma cousine « fatten » .

A mes chères fraires : « Badri, Ammar, Majed, Khalile »

A mes amis : « Abdelmoumen, khairoeddine , seddik, houssem, , Amine, Adel, Hamza, houssem, azzedine, amine , rachide, issam, adem, walide, fakhri »

A mes amies : « Assoum, Ibtissem, nadjat , nardjes, meriem, meriem, majda, saida , khadija , imen, racha, wahiba, nora, sara »

Mon cousins, Mes cousines

A tous ceux qui ont partagés avec moi mes années d'études et les meilleurs moments de ma vie.

Charif

Résumé

Les méthodes de détection des contours sont nombreuses; toutes présentent des avantages mais ne donnent pas entière satisfaction. Le contour actif ou model déformable est l'un des méthodes qui prouve son efficacité. Un contour actif est un ensemble de points qu'on va tenter de déplacer pour leur faire épouser une forme. Son principe est de déplacer les points pour les rapprocher des zones de fort gradient tout en conservant des caractéristiques comme la courbure du contour ou la répartition des points sur le contour ou d'autres contraintes liées à la disposition des points.

Dans ce mémoire nous proposons un système d'extraction de paramètres des anomalies appliqué sur les images mammographiques, Dans un premier temps, nous commençons notre travail avec une phase de prétraitement dans la suite on cherche à détecter les contours de la masse avec un contour actif génétique, qui nous permet dans la suite l'extraction des paramètres de contour élu comme la forme et le type de contour, Dans un deuxième temps, on examine la masse en tenant compte ses caractéristiques, et on calcule l'intensité moyenne et la taille de la masse, les caractéristique détectés nous permet dans nos future travaux à la classification et la décision automatique.

Mots-clefs contours actifs, Prétraitement d'images, segmentation, algorithme génétique, mammographie.

Sommaire

INTRODUCTION GENERALE.....	1
Chapitre I-----	
I.1 Introduction	2
I.2 La Perception & Représentation D'images.....	3
I.3 Historique	4
I.4 L'Image Numérique	4
I.5 Caractéristiques de l'Image Numérique.....	6
I.5.1 Pixel.....	6
I.5.2 Dimension.....	7
I.5.3 Résolution.....	7
I.5.4 Histogramme.....	8
I.5.5 La Luminosité.....	9
I.5.6Le Contraste.....	9
I.5.7Types d'images.....	10
I.5.7.1 Images A Niveaux De Gri	10
I.5.7.2 Images Couleurs.....	11
I.6 Représentation Des Couleurs	11
I.6.1 Synthèse Additive.....	12

I.6.2 Synthèse Soustractive (C.M.J).....	12
I.7 Les Systèmes de Traitement D'images.....	13
I.7.1 Acquisition Des Données Images.....	13
I.7.2 Pré-traitement et Post-Traitement.....	13
I.7.3 Traitement Numérique Des Images.....	14
I.7.3.1 La Convolution.....	14
I.7.3.2 La Segmentation.....	14
I.7.4 Visualisation.....	15
I.8 La Détection de Contours	15
I.8.1 Méthodes Dérivatives.....	17
I.8.1.1 Détection Des Contours Par Dérivée Première.....	17
I.8.1.2 Détection Des Contours Par Dérivée Seconde.....	18
I.8.2 Approche Par Filtrage Optimal.....	19
I.8.3 Les Contours Actifs	20
I.8.3.1 Principe Et Fonctionnement Des Contours Actifs.....	20
I.8.3.2 Domaines D'application	21
I.8.3.3 Les Types De Contours Actifs	22
I.8.3.4 Itérations De L'algorithme.....	22
I.8.3.5 Définition De L'énergie D'un Snake.....	23
A. Énergie Externe.....	23
B. Énergie Interne.....	24

C.Energie De Contexte.....	25
I.8.3.6 Algorithme “Greedy“.....	25
I.9 Conclusion.....	27
Chapitre II-----	
II.1 Introduction.....	28
II.2 Historique	28
II.3 Analogie Avec La Biologie	29
II.3.1 Les Bases des Algorithmes Génétique.....	29
II.3.2 Les Outils Issus De La Biologie.....	30
II.3.2.1 La sélection.....	30
II.3.2.2 Les Croisement Ou Recombinaison.....	31
II.3.2.3 Les mutations.....	31
II.4 Le Principe des Algorithmes Génétiques	32
II.4.1 Le Codage De La Population.....	33
II.4.1.1 Les Types De Codage.....	33
II.4.2 La Fonction D'évaluation.....	35
II.4.3 La Sélection.....	36
II.4.3.1 L'opérateur de sélection.....	36
II.4.3.2 Les méthodes de sélection	36
A-La Loterie Biaisée.....	38
B -La Méthode Elitiste.....	38
C-La Sélection Par K-Tournois.....	38
D-La Sélection Universelle Stochastique.....	38
II.4.4 Le Croisement (Cross Over).....	39
II.4.4.1 L'opérateur de croisement.....	39
II.4.4.2 Les Types De Croisement.....	41

A-Croisement En Un Point.....	41
B -Croisement Multi Point.....	42
C- Croisement Uniforme.....	42
II.4.5La Mutation	42
II.4.5.1 L'opérateur De Mutation	43
II.4.6Création De La Génération Suivante Et Critères D'arrêt.....	45
II.5 Les Apporte Des AGs Par Rapport Aux Autres Paradigmes	45
II.6Les Domaines D'utilisation Des Algorithmes Génétiques.....	46
II.6.1 Applications Dans La Recherche	46
II.6.2 Applications industrielles.....	47
II.6. 3 Informatique Décisionnelle.....	47
II.7Conclusion.....	48
Chapitre III-----	
III.1 Introduction.....	49
II.2 Imagerie Et Algorithmes Génétiques	49
III.3 Les Contours Actifs Génétique	50
III.3.1 Origine.....	50
III.3.2 Implémentation Des Contours Actifs Génétiques.....	50
III.3.2.1Codage Des Chromosomes.....	51
A- Coordonnées Cartésiennes.....	51
B- Coordonnées Polaires.....	53
III.3.2.2Sélection.....	54
III.3.2.3Croisement.....	55
III.3.2.4Mutation.....	56
III.4Conclusion	60

Chapitre VI	
VI.1 Introduction	61
VI.2 Environnement De Développement	61
VI.2.1 Besoins Matériels	61
VI.2.2 Besoins Logiciels	61
VI.2.3 La Version D'Eclipse.....	63
VI.2.4 Le Format DICOM	63
VI.3 Conception	64
VI.3.1 Prétraitement de L'image.....	66
A- Sélection du Segment	66
B- Calcule du gradient	66
C- Changement d'échelle	66
D- Définition du centre du segment	66
E- Filtrage	66
F- Binarisation	66
G- Calcule du Gradient après binarisation	67
VI.3.2 Algorithme Génétique	67
VI.3.2.1 Paramétrage	67
A- Réglage des paramètres de l'algorithme génétique.....	67
B- Réglage des Coefficients des Energies.....	67
C- VI.3.2.2 Initialisation De La population Initiale (Codage).....	68
VI.3.2.3 Evaluation.....	69
VI.3.2.4 Sélection.....	69
VI.3.2.5 Croisement	69
VI.3.2.6 Mutation	70

VI.3.2.7 Critère d'arrêt	71
VI.3.3 Extraction des paramètres de la masse.....	71
VI.3.3.1 Définition des Caractéristique des masses selon BI-RADS.....	72
VI.3.3.2 L'extraction d'informations.....	72
A- La forme de la masse.....	72
B- Type de contours de la masse.....	73.
C- La taille de la masse	73
D- L'intensité de la masse.....	73
VI.4 Réalisation.....	73
VI.4.2 Prétraitement.....	74
VI.4.3 Détection.....	75
VI.4.3.1 Le Déroulement De L'algorithme.....	75
VI.4.3.2 Résultat De L'extraction Des Paramètres.....	77
VI.5 étude de cas.....	78
VI.5.1 La densité de la masse.....	78
VI.5.2 type du contour.....	79
VI.5.3 La Forme de Contours de la Masse.....	79
VI.6 Conclusion & perspective	81

CONCLUSION GENERALE	82

TABLES DES FIGURS

<i>Figure 1.1 : La représentation d'une image</i>	5
<i>Figure 1.2 : Image vectorielle VS image bitmap</i>	6
<i>Figure 1.3: Dimension d'une image</i>	7
<i>Figure 1.4: résolution d'une image</i>	8
<i>Figure 1.5: l'histogramme d'une image couleur (R.V.B)</i>	8
<i>Figure 1.6: la luminance d'une image</i>	9
<i>Figure 1.7: Le contraste d'une image</i>	10
<i>Figure 1.8: Image en niveau de gris</i>	10
<i>Figure 1.9: Image couleur,</i>	11
<i>Figure 1.10: la représentation de la synthèse additive</i>	12
<i>Figure 1.11 : la représentation de la synthèse soustractive</i>	13
<i>Figure 1.12 : schéma du traitement d'image</i>	13
<i>Figure 1.13 : Traitement numérique</i>	14
<i>Figure 1.14 : Exemple d'un contour</i>	15
<i>Figure 1.15 : Les différents types de contours</i>	16
<i>Figure 1.16: Les Méthodes De Détection de contours</i>	16
<i>Figure 1.17: Masques de Gradient (horizontal et vertical)</i>	17
<i>Figure 1.18: Masques de Roberts (horizontal et vertical)</i>	18
<i>Figure 1.19: Masques de Prewitt (horizontal et vertical)</i>	18
<i>Figure 1.20: Masques de sobel (horizontale, verticale, obliques)</i>	18

<i>Figure 1.21 : Principe des contours actifs</i>	21
<i>Figure 1.22 : Types des contours actifs</i>	22
<i>Figure 2.1 : Chromosome humain</i>	30
<i>Figure 2.2 : Croisement humain</i>	31
<i>Figure 2.3 : Mutation humain</i>	32
<i>Figure 2.4 : les étapes d'un algorithme génétique</i>	32
<i>Figure 2.5 :les niveaux d'organisation d'un algorithme génétique</i>	33
<i>Figure 2.6 : Illustration schématique du codage des variables réelle</i>	34
<i>Figure 2.7 : le codage en gray</i>	35
<i>Figure 2.8 :La Fonction D'évaluation</i>	35
<i>Figure 2.9 :La Sélection</i>	36
<i>Figure 2.10 :la méthode de sélection de la loterie biaisée</i>	37
<i>Figure 2.11 :Le Croisement (Cross Over)</i>	39
<i>Figure 2.12 : L'opérateur de croisement</i>	40
<i>Figure 2.13 : le croisement des chromosomes</i>	41
<i>Figure 2.14: croisement avec un point</i>	41
<i>Figure 2.15: croisement multi point</i>	42
<i>Figure 2.16:croisement uniforme</i>	42
<i>Figure 2.17 : L'opérateur De Mutation</i>	43
<i>Figure 3.1 : Codage du snake en coordonnées cartésiennes</i>	51
<i>Figure 3.2 :Initialisation des points en coordonnées cartésiennes sans réordonner les points</i>	52
<i>Figure3.3 : Initialisation des points en coordonnées cartésiennes en réordonnant les points</i>	52
<i>Figure3.4 : les tests de la Sélection par k-tournoi</i>	55
<i>Figure3.5 :comparaison entre les défèrent types de croisement</i>	56

<i>Figure 3.6: les taux de mutation (moyenne sur 25 exécutions)</i>	57
<i>Figure 3.7: Influence de la mutation, zoom sur les générations 150 à 200 (une exécution)</i> ...	58
<i>Figure 3.8 : uniformisation de la population</i>	59
<i>Figure 4.1 : démarrage d'eclipse</i>	63
<i>Figure 4.2 : le diagramme de conception</i>	65
<i>Figure (4.3) : le filtre moyeneur</i>	66
<i>Figure 4.4 : le filtre sobel</i>	67
<i>Figure (4.5) : les coordonnées polaires</i>	68
<i>Figure (4.6) : le tableau de codage des points d'un contour</i>	68
<i>Figure (4.7) : croisement de deux contours en 2 points de coupure P_1 et P_2</i>	70
<i>Figure (4.8) : représentation de la phase de mutation dans CAG</i>	71
<i>Figure 4.9. Les formes des masses</i>	71
<i>Figure 4.10. Les types des contours des masses</i>	72
<i>Figure (4.11): sélection du segment et son centre</i>	73
<i>Figure (4.12): sélection du segment et son centre</i>	74
<i>Figure (4.13): détection du contour</i>	75
<i>Figure (4.14): le déroulement de l'algorithme génétique</i>	76
<i>Figure (4.15): Résultat finale</i>	77
<i>Figure (4.16): la détection sur densités différentes</i>	78
<i>Figure (4.17): contour effacé VS claire</i>	79
<i>Figure (4.18) : les différentes formes du contour</i>	80

LISTE DES TABLEAUX

Tableau 2.1 : voyageur de commerce	46
Tableau (4.1) :Résultat du teste.....	78
Tableau (4.2) :les énergies du contour.....	79
Tableau (4. 3) :les types de forme.....	80

INTRODUCTION GENERALE

L'augmentation de la puissance des calculateurs dans les dernières décennies a permis le développement du traitement d'images numérique, ce traitement qui nécessite un très grand nombre d'opérations mathématiques. Les puissances des processeurs actuels permettent d'effectuer du traitement complexe en temps réel. Les logiciels utilisés pour traiter les photos numériques sont une preuve à cette révolution, dans le domaine du traitement d'images, on distingue deux grandes axes principales : la synthèse d'images et l'analyse d'images, ces deux axes sont étroitement liés. Par exemple, la reconstruction 3D en imagerie médicale utilise les techniques de synthèse pour reconstituer des formes à partir des données fournies par les coupes des scanners ou des balayages ultrasoniques.

L'analyse d'images est le domaine qui nous préoccupe ici particulièrement, elle présente un intérêt dans les domaines quand nous avons besoin de la recherche automatique d'informations dans des images.

Le domaine qui nous intéresse est celui de l'imagerie médicale qui constitue aujourd'hui un outil extraordinaire pour l'aide au diagnostic pour le médecin. Les modalités d'acquisition d'images sont nombreuses et la quantité d'images manipulées s'accroît dans le domaine médical. La radiographie par rayon X constitue la part la plus importante du marché médical, son principe consiste à impressionner sur un film radiographique les différences de densité d'un organe. Le film sera plus ou moins noirci selon l'organe radiographié. L'os apparaît blanc, les tissus mous seront dans différents tons de gris et l'air sera noir.

La mammographie est un examen radiographique du sein, effectué chez la femme pour le dépistage du cancer du sein. Toutefois, la mammographie constitue la modalité principale pour le diagnostic du cancer du sein. L'image mammographique possède également l'avantage d'être très pratique et simple d'utilisation. L'analyse des images mammographique permet de fournir au médecin des informations quantitatives fiables en 2D, ces informations complètent l'aspect visuel et qualitatif de l'examen. Du fait de la nature très bruitée des images mammographique, la délimitation visuelle précise des lésions mammaires est un problème réel pour le médecin. Les contours des différentes régions pathologiques sont rarement bien délimités, d'autres examens complémentaires sont alors nécessaires pour compléter le diagnostic, par exemple l'échographie et l'imagerie par résonance magnétique (IRM).

L'objet du travail décrit dans ce mémoire s'inscrit dans le domaine de la détection des contours d'images et l'extraction de paramètres. Le travail que nous avons effectué porte sur la détection des contours flous bruités ou texturés telles que les données trouvées sur l'organe mammaire. La nature des données rend la détection des contours très difficile avec les approches de détection classiques.

Le premier chapitre de ce mémoire est consacré aux méthodes de détection des contours et la vision artificielle; où nous rappelons les différentes techniques classiques de la détection de contours, et un état de l'art sur les contours actifs ; Ainsi le premier chapitre permet de poser les éléments mathématiques qui nous serviront tout au long du développement. Dans le deuxième chapitre nous présentons un état de l'art sur la méthode des algorithmes génétiques, Le troisième chapitre est consacré à la présentation d'une méthode des contours actifs basée sur les algorithmes génétiques [RO03], cette méthode que nous avons appliquée pour la recherche des contours d'une éventuelle anomalie sur la mammographie, et pour l'extraction de paramètres liés à la forme et le type de contour de la masse.

Dans le quatrième chapitre nous présentons la conception de notre travail, ainsi que l'application testée sur des images DICOM. une application pour l'aide au diagnostic appliquée sur les images mammographiques, cette application permet l'extraction de paramètres de l'anomalie.

Enfin, la conclusion générale présente une synthèse de ce travail et les perspectives qui en découlent.

**Chapitre 01 : Traitement d'image & vision
artificielle**

I.1 Introduction

Avec la parole, l'image constitue l'un des moyens les plus importants qu'utilise l'homme pour communiquer avec autrui. C'est un moyen de communication universel dont la richesse du contenu permet aux êtres humains de tout âge et de toute culture de se comprendre. Chacun peut analyser l'image à sa manière, pour en dégager une impression et d'en extraire des informations précises. De ce fait, le domaine d'imagerie est l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes.

La vision par ordinateur désigne la compréhension d'une scène ou d'un phénomène à partir d'informations « image », liant intimement perception, comportement et contrôle.

Le traitement d'image et la vision artificielle sont des disciplines relativement jeunes (~années 60) et qui évoluent rapidement. Elles sont en plein expansion et donnent lieu chaque année à une profusion de travaux, académiques, technologiques, industriels.

Ce chapitre abordera les techniques d'acquisition, de traitement, d'analyse et d'interprétation des images numériques dans l'objectif de construire un système de vision artificielle.

I.2 La Perception & Représentation D'images

La *perception* d'une image s'articule autour des caractéristiques du système visuel humain. Ainsi, il apparaît que l'œil est sensible à certaines fréquences du spectre électromagnétique; ces fréquences représentent la lumière en général. Dans certains cas, le rayonnement électromagnétique à observer se situe en dehors de la plage des fréquences visibles (les rayons X par exemple). On doit alors recourir à un révélateur pour transposer le signal à observer dans une plage visible.

La *représentation* d'une image. Il s'agit de représenter une entité physique sous une forme électrique ou une forme informatique. La représentation joue un rôle essentiel dans une chaîne de traitement car elle conditionne la capacité de stockage nécessaire ainsi que la mise en œuvre. En outre se pose régulièrement la question du changement de représentation. À titre d'illustration, une séquence vidéo au format PAL nécessite un débit de 199Mb/s sans compression. Le changement de représentation que représente la compression ramène ce débit à des valeurs de l'ordre de 4 à 9 Mb/s pour une qualité similaire. [CF88]

I.3 Historique

Dans les années 1920 : Le traitement d'images commence à être étudié pour la transmission d'images par le câble sous-marin allant de New York à Londres. Harry G. Bartholomew et Maynard D. Mc Farlane effectuent la première numérisation d'image avec compression de données pour envoyer des fax de Londres à New York.

Les années 1960 : Le besoin de traiter les images est rapidement devenu une évidence en plusieurs domaines : physique des particules, en médecine, chimie, et les chercheurs commençaient à s'intéresser à l'automatisation de l'interprétation des images, ainsi qu'à la reconnaissance des formes que ce soit pour déchiffrer un texte dactylographié ou manuscrit, pour compter des chromosomes, reconnaître une tumeur, un char ou un avion de guerre.

Les années 1980 : La vision par ordinateur s'est développée grâce à David Marr (1945-1980). Ce neurophysiologiste et mathématicien anglais a été un des premiers à définir les bases formelles de la vision par ordinateur en intégrant des résultats issus de la psychologie et de l'intelligence artificielle.

De nos jours, le traitement d'images c'est répandu et le grand public a pris part grâce à la capacité croissante des ordinateurs et autres appareils portables, capables de traiter des images presque instantanément. Néanmoins, même les ordinateurs les plus récents ne sont pas encore assez performants pour résoudre certains problèmes, en particulier en imagerie médicale. Les quantités d'information à stocker et à traiter restent problématiques en revanche, beaucoup d'applications sont déjà courantes comme les traitements sur les images des appareils photo numériques que chacun peut faire sur son ordinateur personnel, ou les effets spéciaux au cinéma pour mélanger des images virtuelles tridimensionnelles à une scène [HK10].

I.4 L'Image Numérique

L'image est un ensemble structuré d'informations qui après affichage sur l'écran, ont une signification pour l'œil humain. Elle peut être décrite sous la forme d'une fonction $I(x,y)$ de brillance analogique continue, définie dans un domaine borné, tel que x et y sont les coordonnées spatiales d'un point de l'image et I est une fonction d'intensité lumineuse et de couleur. Sous cet aspect, l'image est inexploitable par la machine, ce qui nécessite sa numérisation [MA87].

Contrairement aux images obtenues à l'aide d'un appareil photo, ou dessinées sur du papier, les images manipulées par un ordinateur sont numériques (représentées par une série de bits).

L'image numérique est l'image dont la surface est divisée en éléments de tailles fixes appelés cellules ou pixels, ayant chacun comme caractéristique un niveau de gris ou de couleurs prélevé à l'emplacement correspondant dans l'image réelle, ou calculé à partir d'une description interne de la scène à représenter [GW97].

La numérisation d'une image est la conversion de celle-ci de son état analogique (distribution continue d'intensités lumineuses dans un plan xOy) en une image numérique représentée par une matrice bidimensionnelle de valeurs numériques $f(x,y)$ où :

x, y : coordonnées cartésiennes d'un point de l'image comme c'est montré

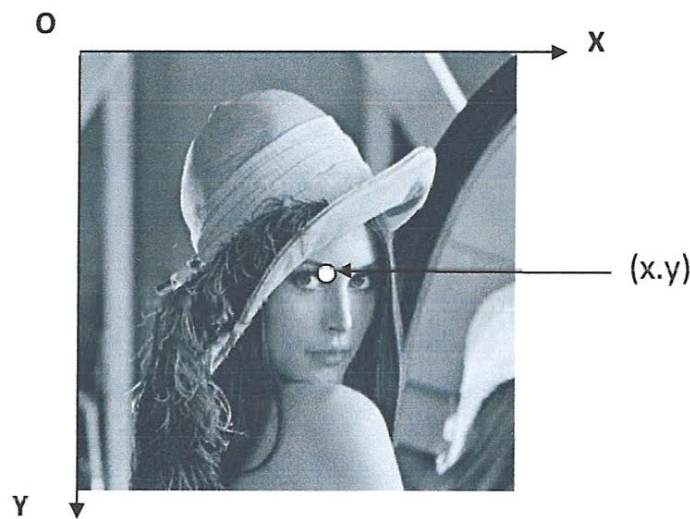


Figure 1.1 : La représentation d'une image.

$f(x, y)$: niveau de gris en ce point

Pour des raisons de commodité de représentation pour l'affichage et l'adressage, les données images sont généralement rangées sous formes de tableau I de n lignes et p colonnes. Chaque élément $I(x, y)$ représente un pixel de l'image et à sa valeur est associé un niveau de gris codé sur m bits (2^m niveaux de gris ; $0 = \text{noir}$; $2^m - 1 = \text{blanc}$). La valeur en chaque point exprime la mesure d'intensité lumineuse perçue par le capteur [JM08].

Il existe 2 sortes d'images numériques : les images *matricielles* sbitmap (image-bit) et les images *vectérielles*. Alors qu'une image vectorielle est décrite à l'aide de courbes et d'équations mathématiques, une image bitmap est constituée de pixels et se réduit donc à une matrice de points.

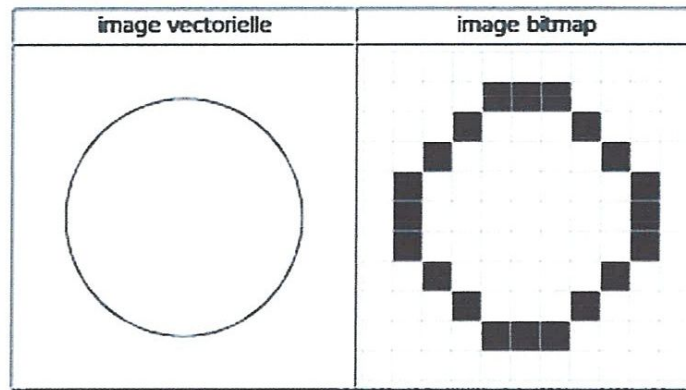


Figure 1.2 : Image vectorielle VS image bitmap.

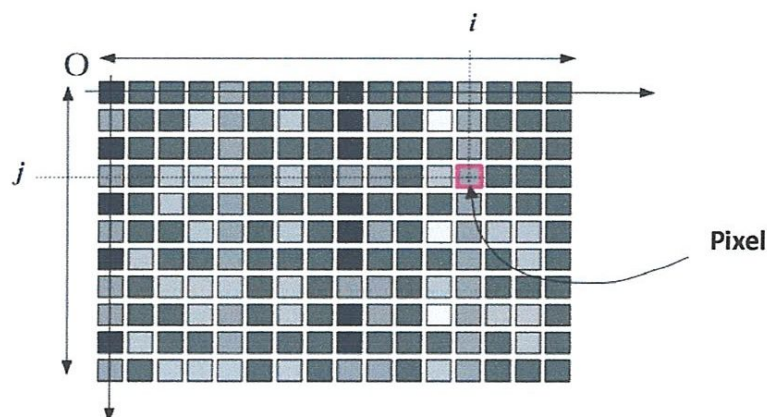
Si les images vectorielles peuvent être manipulées avec beaucoup de facilité, les modifications de taille, par exemple, apportées à une image bitmap ne sont pas sans incidence [KC99].

1.5 Caractéristiques de l'Image Numérique

L'image est un ensemble structuré d'informations caractérisé par les paramètres suivants:

1.5.1 Pixel

Contraction de l'expression anglaise " picture-element ": éléments d'image, le pixel est le plus petit point de l'image, c'est une entité calculable qui peut recevoir une structure et une quantification. Si le bit est la plus petite unité d'information que peut traiter un ordinateur, le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels d'affichage ou d'impression.



La quantité d'information que véhicule chaque pixel donne des nuances entre images monochromes et images couleurs. Dans le cas d'une image monochrome, chaque pixel est codé sur un octet, et la taille mémoire nécessaire pour afficher une telle image est directement liée à la taille de l'image.

Dans une image couleur (R.V.B.), un pixel peut être représenté sur trois octets : un octet pour chacune des couleurs : rouge (R), vert (V) et bleu (B)[JM08].

I.5.2 Dimension

C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image [JM08].

$$Dim(dimension) = (h)hauteur * (l)largeur \quad (I.1)$$



Figure 1.3: Dimension d'une image.

I.5.3 Résolution

C'est la clarté ou la finesse de détails atteinte par un moniteur ou une imprimante dans la production d'images. Sur les moniteurs d'ordinateurs, La résolution d'une image représente le nombre de pixels par pouce, **PPI** (en anglais Pixel Per Inch) en ce qui concerne les supports numériques et par le nombre de points par pouce **DPI** (en anglais Dot Per Inch) en ce qui concerne l'impression papier. Ainsi, une image de 3x3 pouces avec une résolution de 300 PPI contiendra 900x900 pixels. Le terme résolution est aussi employé pour désigner le nombre de pixels contenus dans une image, on parle alors du nombre de pixels par ligne et du nombre de pixel par colonne[KH08].

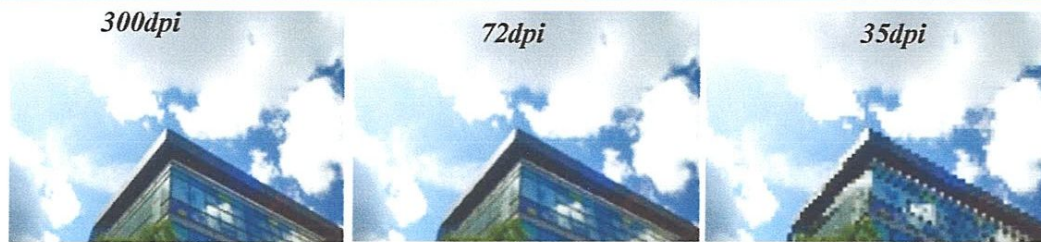


Figure 1.4: résolution d'une image.

I.5.4 Histogramme

L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleur) dans l'image. Pour diminuer l'erreur de quantification, pour comparer deux images obtenues sous des éclairages différents, ou encore pour mesurer certaines propriétés sur une image, on modifie souvent l'histogramme correspondant [GW97].

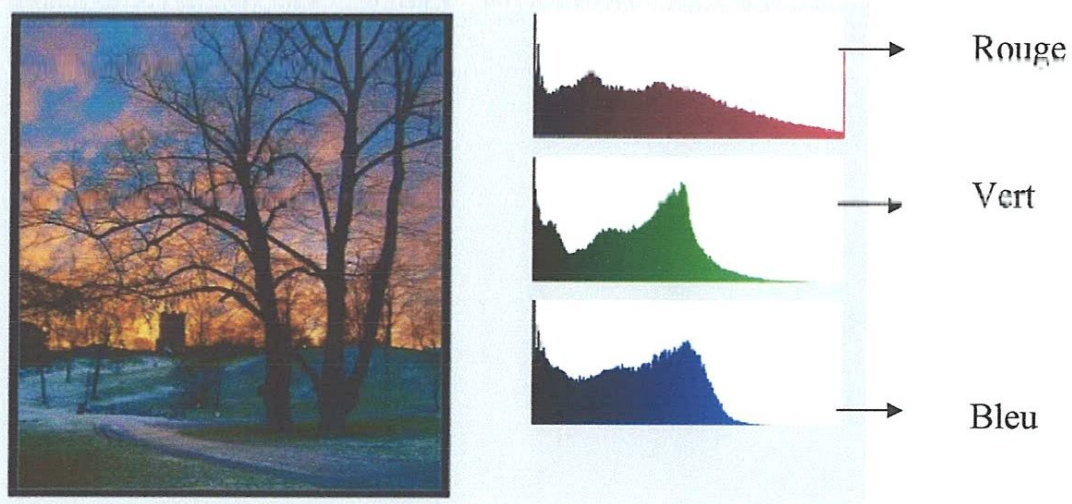


Figure 1.5: l'histogramme d'une image couleur (R.V.B).

Il permet de donner un grand nombre d'information sur la distribution des niveaux de gris (couleur) et de voir entre quelles bornes est répartie la majorité des niveaux de gris (couleur) dans les cas d'une image trop claire ou d'une image trop foncée. Il peut être utilisé pour améliorer la qualité d'une image (Rehaussement d'image) en introduisant quelques modifications, pour pouvoir extraire les informations utiles de celle-ci.

I.5.5 La Luminosité

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface, pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet.

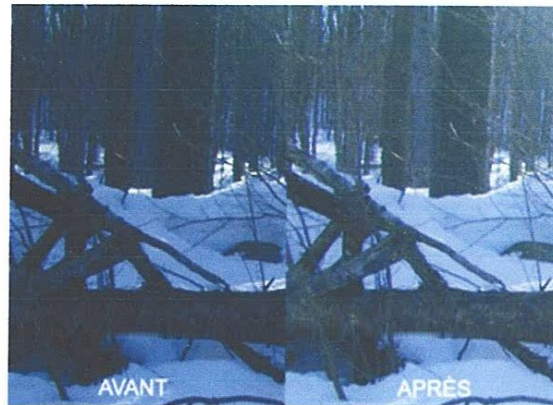


Figure 1.6: la luminance d'une image.

Une bonne luminance se caractérise par :

1. Des images lumineuses (brillantes)
2. Un bon contraste : il faut éviter les images où la gamme de contraste tend vers le blanc ou le noir; ces images entraînent des pertes de détails dans les zones sombres ou lumineuses.
3. L'absence de parasites [JM08].

I.5.6Le Contraste

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images[ST06]. Si L_1 et L_2 sont les degrés de luminosité respectivement de deux zones voisines A_1 et A_2 d'une image, le contraste C est défini par le rapport :

$$C = \frac{L_1 - L_2}{L_1 + L_2} \quad (I.2)$$

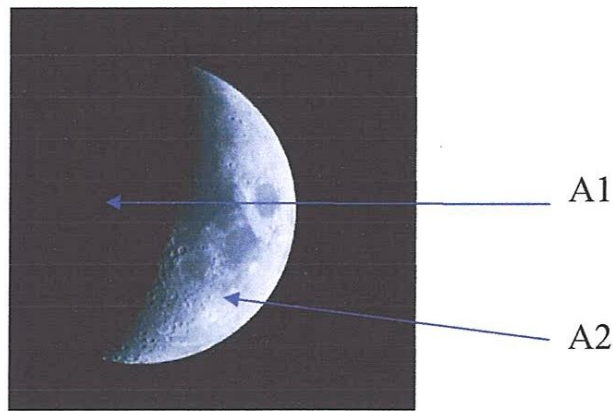


Figure 1.7: Le contraste d'une image.

I.5.7 Types d'images

I.5.7.1 images A Niveaux De Gris

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires. Donc pour représenter les images à niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise par exemple entre 0 et 255. Chaque pixel n'est donc plus représenté par un bit, mais par un octet. Pour cela, il faut que le matériel utilisé pour afficher l'image soit capable de produire les différents niveaux de gris correspondant.



Figure 1.8: Image en niveau de gris.

Le nombre de niveaux de gris dépend du nombre de bits utilisés pour décrire la "couleur" de chaque pixel de l'image. Plus ce nombre est important, plus les niveaux possibles sont nombreux [JM08].

I.5.7.2 Images Couleurs

Même s'il est parfois utile de pouvoir représenter des images en noir et blanc, les applications multimédias utilisent le plus souvent des images en couleurs. La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités. En effet, il faut tout d'abord choisir un modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange de rouge, vert et bleu (modèle R.V.B par exemple.)[JM08].



Figure 1.9: Image couleur.

I.6 Représentation des Couleurs

Afin de pouvoir manipuler correctement des couleurs et échanger des informations colorimétriques il est nécessaire de disposer de moyens permettant de les catégoriser et de les choisir. Ainsi, il n'est pas rare d'avoir à choisir la couleur d'un produit avant même que celui-ci ne soit fabriqué. Dans ce cas, une palette de couleurs est présentée, dans laquelle la couleur convenant le mieux au besoin est choisie. La plupart du temps le produit (véhicule, bâtiment, etc.) possède une couleur qui correspond à celle choisie.

En informatique, de la même façon, il est essentiel de disposer d'un moyen de choisir une couleur parmi toutes celles utilisables. Or la gamme de couleurs possibles est très vaste et la chaîne de traitement de l'image passe par différents périphériques : par exemple un numériseur (scanner), puis un logiciel de retouche d'image et enfin une imprimante. Il est donc nécessaire de pouvoir représenter de façon fiable la couleur afin de s'assurer de la

cohérence entre ces différents périphériques. On appelle ainsi «**espace de couleurs**» la représentation mathématique d'un ensemble de couleurs. Il en existe plusieurs, parmi lesquels les plus connus sont :

- Le codage **RGB** (*Rouge, Vert, Bleu*, en anglais **RGB**, *Red, Green, Blue*).
- Le codage **TSL** (*Teinte, Saturation, Luminance*, en anglais **HSL**, *Hue, Saturation, Luminance*).
- Le codage **CMYK**.
- Le codage **CIE**.
- Le codage **YUV**.
- Le codage **YIQ**.

Il existe deux types de synthèse de couleur:

I.6.1 Synthèse Additive

Elle consiste à utiliser 24 bits pour chaque point de l'image. Huit bits sont employés pour décrire la composante rouge (R), huit pour le vert (V) et huit pour le bleu (B). Il est ainsi possible de représenter environ 16,7 millions de couleurs différentes simultanément. Cela est cependant théorique, car aucun écran n'est capable d'afficher 16 millions de points. Dans la plus haute résolution (1600 x 1200), l'écran n'affiche que 1 920 000 points. Par ailleurs, l'œil humain n'est pas capable de distinguer autant de couleurs [JM08].

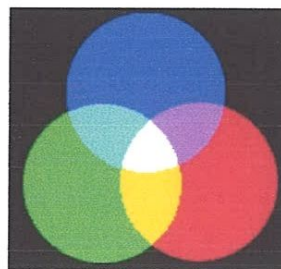


Figure 1.10: la représentation de la synthèse additive.

I.6.2 Synthèse Soustractive (C.M.J)

La synthèse soustractive permet de restituer une couleur par soustraction, à partir d'une source de lumière blanche, avec des filtres correspondant aux couleurs complémentaires : jaune,

magenta, et cyan. L'ajout de ces trois couleurs donne du noir et leur absence produit du blanc [DM11].

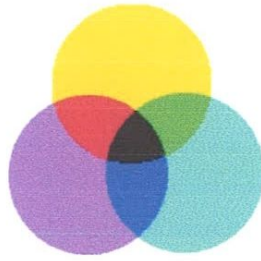


Figure 1.11 : la représentation de la synthèse soustractive

I.7 Les Systèmes de Traitement D'images

Un système de traitement numérique d'images est composé de

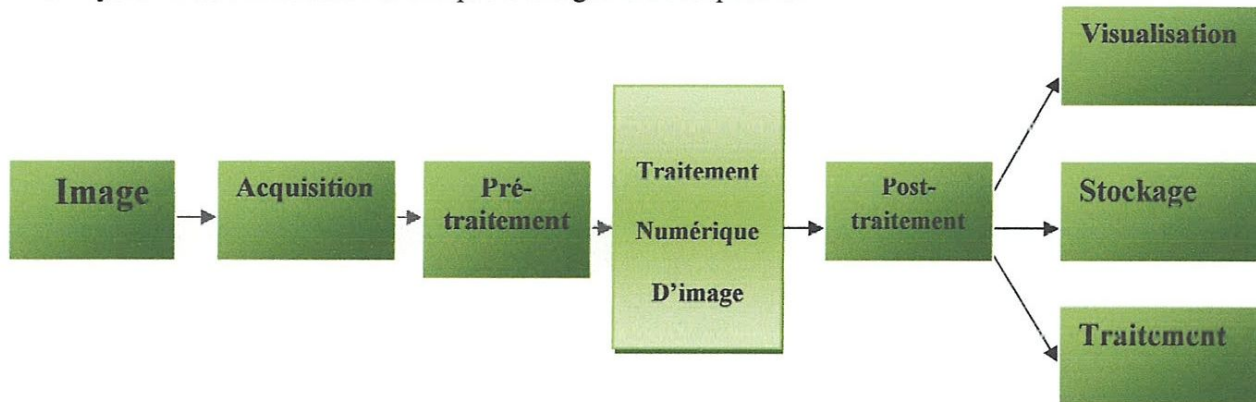


Figure 1.12 : schéma du traitement d'image.

I.7.1 Acquisition Des Données Images

L'acquisition d'images constitue un des maillons essentiels de toute chaîne de conception et de production d'images. Pour pouvoir manipuler une image sur un système informatique, il est avant tout nécessaire de lui faire subir une transformation qui la rendra lisible et manipulable par ce système. Le passage de cet objet externe (l'image d'origine) à sa représentation interne (dans l'unité de traitement) se fait grâce à une procédure de numérisation.

I.7.2 Pré-traitement et Post-Traitement

Les images numériques telles qu'elles sont acquises, sont très souvent inexploitable pour le traitement d'images. Elles contiennent des signaux bruités (pas assez de contraste ou un

éclairage insuffisant). Pour remédier à cela, différents prétraitements pour l'amélioration, la correction ou le rehaussement de l'image sont effectués

Le prétraitement concerne les images acquises tandis que le post-traitement concerne les images traitées. Dans les deux cas, les images ne peuvent échapper aux effets de dégradations dus essentiellement aux phénomènes physiques tels que :

- La diffraction (déviations) du système optique.
- Le flou dû au mouvement de l'image durant son acquisition.

Pour pallier à ces dégradations on utilise en général le filtrage.

I.7.3 Traitement Numérique Des Images

Dans cette étape on peut citer comme titre d'exemples les traitements suivants :

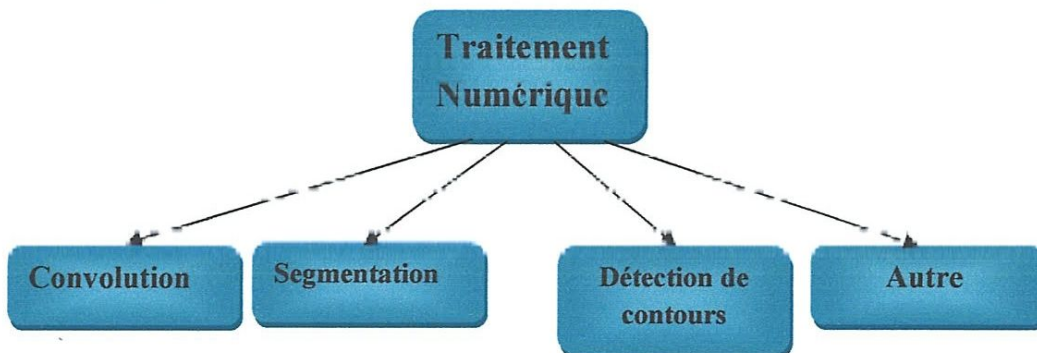


Figure 1.13 : Traitement numérique.

I.7.3.1 La Convolution

La convolution est le remplacement de la valeur d'un pixel par une combinaison linéaire de ses voisins. Elle consiste à faire balayer une fenêtre (masque) sur l'ensemble des points de l'image[KC99].

I.7.3.2 La Segmentation

On est conduit pour analyser une image à la partager en un certain nombre de domaines connexes : ce processus joue un rôle de plus en plus important en traitement numérique des images.

Deux points appartiennent à un même domaine s'ils possèdent une même propriété caractéristique P et si on peut les joindre par une chaîne continue de points voisins ayant cette propriété P. Une segmentation est donc définie par le type de propriété et par la définition de voisinage [KC99].

I.7.4 Visualisation

Tout système de traitement d'image est doté d'un dispositif de visualisation qui permet l'affichage des images. L'utilisation de différents types de reconstituteurs permet de transformer le signal numérique qu'est la matrice image en un signal analogique visible par l'œil de l'observateur. Pour cela, différents types de supports peuvent être employés : moniteur vidéo, clichés photographiques, impression sur papier. Dans tous les cas et pour chaque échantillon de l'image numérique, on recrée un nouvel élément d'image ou un nouveau pixel dont on choisit la forme de façon à reconstituer une image analogique qui soit la plus proche possible de l'image avant numérisation compte tenu des erreurs introduites lors de l'acquisition, de la numérisation et de la transmission [KC99].

I.8 La Détection de Contours

Un contour est un ensemble de pixels formant une frontière entre deux ou plusieurs régions voisines. L'épaisseur du contour est d'un pixel ou plus (un pixel dans le cas idéal). Une frontière existe s'il a eu discontinuité du niveau de gris entre deux pixels adjacents (une brusque variation du niveau de gris ou de couleur).

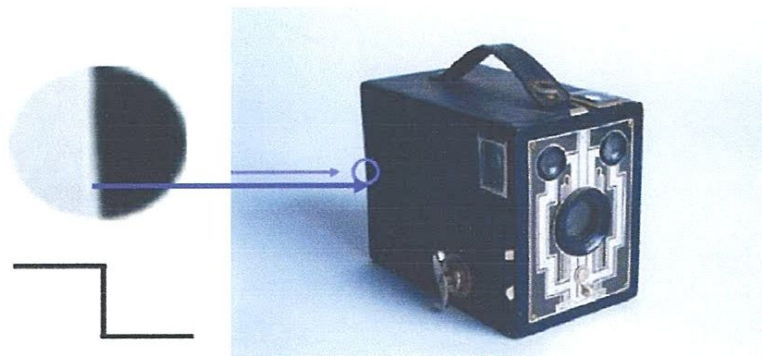


Figure 1.14 : Exemple d'un contour.

Il existe 3 types de contours, la figure (1.15) montre les différents types de contours :

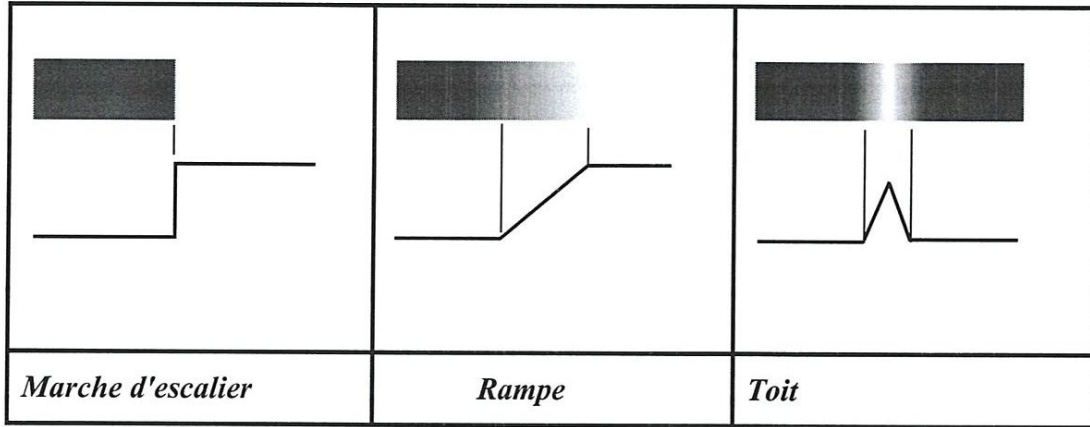


Figure 1.15 : Les différents types de contours.

La détection de contours est l'une des opérations de base de tout système de vision artificielle (reconnaissance de caractères, détection de défauts sur des pièces industrielles...). C'est un processus qui essaie de capturer les propriétés significatives des objets dans l'image. Ces propriétés incluent des discontinuités dans les caractéristiques photométriques géométriques et physiques des objets.

La détection des contours d'une image réduit de manière significative la quantité de données et élimine les informations qu'on peut juger moins pertinentes, tout en préservant les propriétés structurelles importantes. qu'elles sont toutes précédées par d'autres méthodes de prétraitement

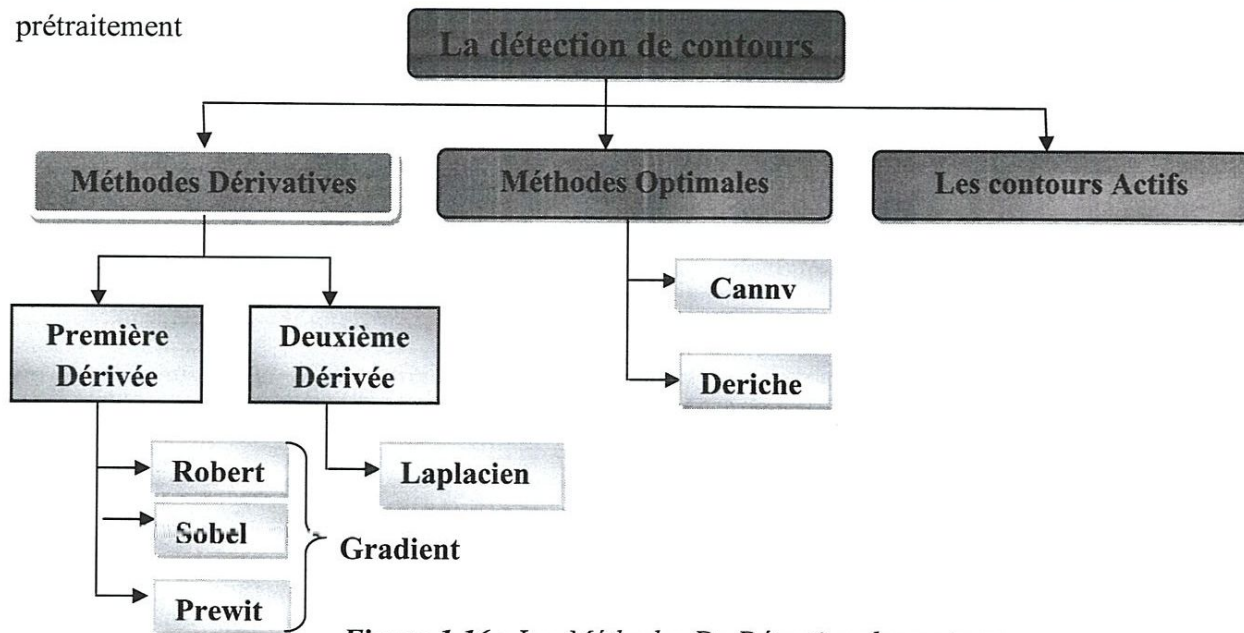


Figure 1.16 : Les Méthodes De Détection de contours.

Il existe un grand nombre de méthodes de détection de contours. Les plus utilisées sont les approches à base de dérivée première (Gradient) et dérivée seconde (Laplacien) et les approches optimales, malheureusement, ces méthodes ne donnent pas directement de bons résultats satisfaisants à cause de leur sensibilité au bruit, c'est pourquoi qu'elles sont toutes précédées par d'autres méthodes de prétraitement. [RR09]

I.8.1 Méthodes Dérivatives

I.8.1.1 Détection Des Contours Par Dérivée Première

Ce sont des filtres de dérivée première (appelés aussi filtres étroits), et l'on cherche alors le maximum de leur réponse[1]. Le gradient est une dérivation au premier ordre, il est donné par la formule :

$$\vec{\nabla}I = \frac{\partial I}{\partial x} \vec{i}_x + \frac{\partial I}{\partial y} \vec{i}_y \quad (1.3)$$

Où \vec{i}_x (resp. \vec{i}_y) Est un vecteur unitaire suivant x (resp. suivant y).

En un pixel d'une image numérique, le gradient est un vecteur caractérisé par son amplitude et sa direction. L'amplitude est directement liée à la quantité de variation locale des niveaux de gris.

La direction du gradient est orthogonale à la frontière qui passe au point considéré.

Appliquée à une image, considérée comme une fonction de deux variables, on peut définir deux dérivées partielles, suivant x (colonnes) et suivant y (lignes). Les masques correspondants figurent sur la figure suivante (1.17):

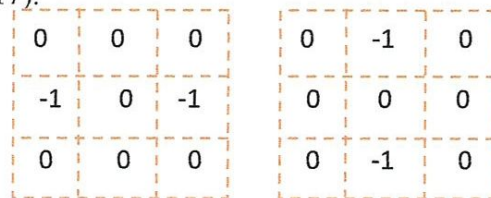


Figure 1.17: Masques de Gradient (horizontal et vertical).

La dérivation accentuant le bruit (pixels parasites de répartition aléatoire), des filtres dérivés plus robustes comme (Robert, Sobel, Prewitt...) qui ont globalement les mêmes propriétés.[1]

❖ Filtre de Roberts

Le détecteur de Roberts permet de calculer le gradient bidimensionnel d'une image de la manière simple et rapide calculé sur 4 points, ce filtre consiste en application du double masque suivant :

$$\begin{array}{|c|c|} \hline 1 & 0 \\ \hline 0 & -1 \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline -1 & 0 \\ \hline \end{array}$$

Figure 1.18: Masques de Roberts (horizontal et vertical).

❖ Filtre de Prewitt

Calculé sur 9 points, ce filtre effectue une moyenne locale sur 3 points en même temps que la dérivation[1]. Il est défini par le double masque suivant (normalisé par un facteur de 1/3) :

$$\begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array}$$

Figure 1.19: Masques de Prewitt (horizontal et vertical).

❖ Filtre de Sobel

Également calculé sur 9 points, ce filtre est une variante du filtre de Prewitt mais permet de privilégier le calcul suivant certaines directions (horizontale, verticale, obliques). [1]

Les filtres sont donc les suivants (figure 1.20) :

$$\begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Figure 1.20: Masques de sobel (horizontale, verticale, obliques).

1.8.1.2 Détection Des Contours Par Dérivée Seconde

L'utilisation des dérivées secondes est plus facile que celle des dérivées premières, au lieu de chercher les maximums de l'intensité du gradient, on cherche les zéros de la dérivée seconde plus précisément le Laplacien qui est une dérivation au deuxième ordre :

$$\Delta I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \quad (\text{I.4})$$

Dans le cas d'une image, il n'existe pas une dérivée seconde unique mais quatre dérivées partielles (selon x^2 , y^2). En pratique, on lève cette ambiguïté en ayant recours à l'opérateur Laplacien qui fait la somme des deux dérivées partielles principales seulement.

L'estimation de la dérivée seconde étant très sensible aux bruits, il convient de filtrer très fortement l'image avant d'en mesurer le Laplacien. Ainsi, afin de limiter les réponses dues au bruit de l'image I , le plus souvent, on fait appel à un filtrage gaussien dont le Laplacien est plus connu sous le nom de "chapeau mexicain". [1]

I.8.2 Approche Par Filtrage Optimal

Un autre type d'approche plus récentes repose sur la définition de critères d'optimalité de la détection de contours; ces critères débouchant sur des filtres de lissage *optimaux*.

Les filtres optimaux sont utilisés largement dans les années quatre-vingt-dix en reconnaissance des formes, en analyse d'image.[1]

Filtre De Canny (1986)

Le filtre de Canny est un filtre impair à réponse impulsionnelle finie (RIF) défini sur l'intervalle $[-M;M]$ ce qui signifie que h est nul en dehors de cet intervalle. La résolution du système est assez complexe (on maximise le produit $\Sigma\Lambda$ sous la contrainte du troisième terme constant). [1]

Les approches précédentes par masques étaient basées sur une modélisation assez simple d'un contour idéal. Pour le calcul du filtre de Canny, une approche analytique plus élaborée est employée. Il s'agit d'une technique de filtrage optimal. Canny a en effet cherché à déterminer de façon analytique en 1986 un filtre à partir de trois critères :

- ✓ Un critère de bonne détection garantissant une réponse forte en sortie du filtre même en présence de faibles contours sur l'image d'entrée.
- ✓ Un critère de bonne localisation du contour.
- ✓ Un critère d'unicité de la réponse permettant d'assurer une seule détection pour un contour et ainsi d'éviter les effets de rebond. [2]

Filtre De Deriche

Deriche a proposé un filtre de lissage dont la dérivée est la solution exacte de l'équation de Canny étendue aux filtres à supports infinis.

Le détecteur de Deriche répond exactement aux mêmes Critères de qualité que celui de Canny mais qui possède une réponse impulsionnelle infinie. [1]

I.8.3 Les Contours Actifs

Les contours actifs tirent leur origine des modèles élastiques, mais la communauté s'accorde à les attribuer à l'équipe « Kass, Witkin et Terzopoulos » qui ont introduit les snakes ou les courbes minimisantes.

Le principe des contours actifs est de faire évoluer un contour fermé initial vers une position d'équilibre, c'est-à-dire en direction des bords de l'objet à détecter. Entre deux itérations, la vitesse de déplacement des points dépend de l'équation d'évaluation du contour.

Et cette équation dépend de son tour aux données présentes dans l'image (intensité, gradient...) et des propriétés de régularité du contour. La position d'équilibre peut être définie comme le minimum d'une fonctionnelle d'énergie dont la dérivée correspond aux forces à appliquer.

Les snakes tiennent leur nom de leur aptitude à se déformer comme des serpents, ils sont devenus un sujet très important pour la communauté du traitement d'images. De très nombreuses équipes s'y sont intéressées de manière plus ou moins approfondie. « Mc Inerney » qui a proposé une vue d'ensemble des modèles déformables en analyse d'images médicales, une revue des méthodes de segmentation d'images médicales proposé par « Pharm et All ». [ROU03]

I.8.3.1.Principe Et Fonctionnement Des Contours Actifs

Un contour actif est un ensemble de points qu'on va tenter de déplacer pour leur faire épouser une forme. Il s'agit d'une technique d'extraction de données utilisée en traitement d'images. L'idée de cette méthode est de déplacer les points pour les rapprocher des zones de fort gradient tout en conservant des caractéristiques comme la courbure du contour ou la répartition des points sur le contour ou d'autres contraintes liées à la disposition des points.

Au démarrage de l'algorithme, le contour est disposé uniformément autour de l'objet à détecter puis il va se rétracter pour en épouser au mieux ses formes. De la même manière, un contour actif

peut aussi se dilater et tenter de remplir une forme, il sera alors situé à l'intérieur de celle-ci au démarrage de l'algorithme.

A chaque itération, l'algorithme va tenter de trouver un meilleur positionnement pour le contour pour minimiser les dérives par rapport aux contraintes utilisées. L'algorithme s'arrêtera lorsqu'il ne sera plus possible d'améliorer le positionnement ou simplement quand le nombre maximum d'itérations aura été atteint. On utilise les notions d'énergies interne et externe pour caractériser respectivement la forme du contour et tous les éléments qui lui sont propres, et le positionnement du contour sur l'image en tenant compte des lignes de gradient. [PS08]

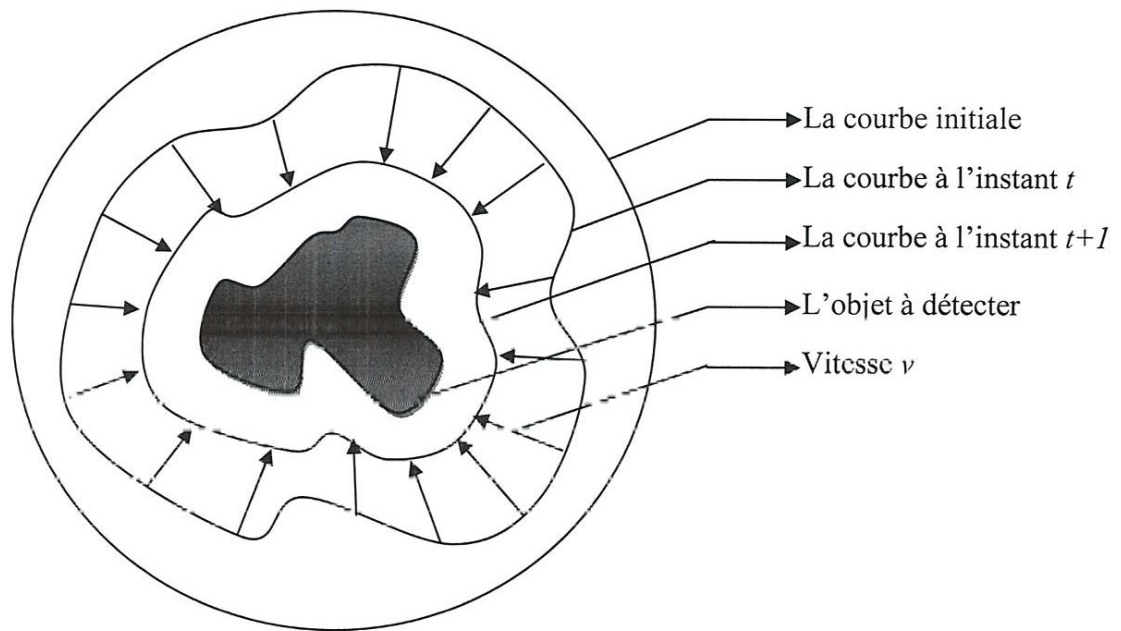


Figure 1.21 : Principe des contours actifs

I.8.3.2 Domaines D'application

Les domaines d'utilisation des contours actifs sont nombreux tant en 2D qu'en 3D tels : la reconnaissance de formes, la simulation, le suivi de scènes, la segmentation d'images.

Du fait que les surfaces intéressantes en imagerie médicale sont en général des formes bien régulières, l'utilisation des modèles déformables est particulièrement intéressante pour déterminer les surfaces de frontières d'organes et de structures. Nous avons commencé à utiliser les contours actifs détecter les contours de la glande mammaire et pour délimiter les zones suspects dans les mammographies. [COU05]

I.8.3.3 Les Types De Contours Actifs

Malgré les difficultés de réglage de la convergence, les contours actifs apportent une solution heureuse à la détection de contour, intermédiaire entre les solutions purement locales (opérateurs de filtrage vus précédemment) et les segmentations globales. Nous pouvons définir trois types de contours actifs déferents utilisés :

1. les contours actifs fermés où $v_0^t = v_{n-1}^t$ (le point de départ est celui d'arrivée).
2. les contours actifs à extrémités libres (tous les points sont susceptibles de bouger).
3. les contours actifs à extrémités fixes où les positions v_0 et v_{n-1} sont fixes dans le temps ainsi que les dérivées premières en ces points (les points de départ et d'arrivée restent fixes). [COU05]

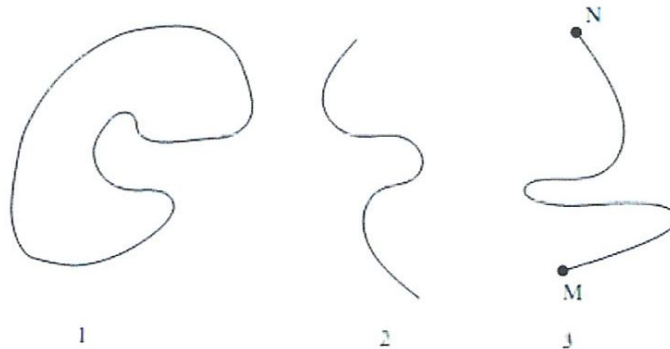


Figure 1.22 : Types des contours actifs.



I.8.3.4 Itérations De L'algorithme

Chaque itération peut se représenter de la manière suivante :

- calcul des énergies interne et externe, caractérisant le contour lui-même et son positionnement sur l'image.
- pour chaque point du contour, détermination d'une nouvelle position, sur laquelle le contour devrait mieux
- minimiser les écarts de contraintes.
- arrangement du contour pour qu'il respecte des contraintes d'écartement entre les points, de régularité de points ...

Note: Les énergies peuvent être calculées dans un espace réduit (l'image discrétisée à chaque point du snake) ou bien dans le domaine image, en considérant une spline reliant tous les pixels du snake.

Cette méthode, bien que plus rigoureuse mathématiquement, n'offre pas de réelles améliorations et apporte plus de temps de calcul qu'autre chose. [PS08]

1.8.3.5 Définition De L'énergie D'un Snake

Deux types d'énergies sont utiles au Snake. Premièrement, puisque son rôle est de délimiter certaines régions d'une image, il existe une énergie liée à l'image. Plus généralement, nous parlerons de l'énergie externe $E_{externe}$ qui guide le contour actif sur l'image. Deuxièmement, la méthode de Kass introduit un terme de régularisation [BOU09].

Ce terme se traduit par une énergie interne $E_{interne}$ imposant des contraintes sur la géométrie lors de la déformation. L'énergie du Snake E_{snake} s'écrit alors :

$$E_{snake}(g(s)) = E_{interne}(g(s)) + E_{externe}(g(s)) \quad (I.5)$$

A. Énergie Externe

La vocation de l'énergie externe est d'attirer le contour actif vers la position correspondant à une caractéristique recherchée. Sa définition est donc fortement dépendante de l'application. Cependant, l'utilisation générale des Snakes et la détection de contour. C'est pourquoi l'énergie la plus utilisée est le gradient de l'image. C'est également la plus simple des cartes de contours que l'on puisse obtenir.

L'énergie externe est également une stratégie d'introduction de contraintes spécifiées par l'utilisateur [BOU09]. On peut représenter cette énergie par le gradient ou l'intensité.

Gradient : Pour la recherche des zones de fort contraste dans l'image, est introduite la fonction :

$$E = - \int_0^1 \|\nabla I(v(s))\|^2 ds \quad (I.6)$$

Où $\nabla I(v(s))$ représente le gradient de l'image $I(v(s))$ [BOU09].

L'Intensité : Cette énergie, au contraire, permet de sélectionner les zones sombres ou claires selon le signe choisi.

$$E = \pm \int_0^1 (I(v(s) - i_0))^2 ds \quad (I.7)$$

Les valeurs i_0 introduit ou non, un certain seuillage. On peut ainsi favoriser la position du contour dans une zone donnée [BOU09].

B. Énergie Interne

L'énergie interne ne dépend pas de l'image ni de la forme à détourner, elle ne dépend que des points du contour. Elle regroupe des notions comme la courbure du contour ou la régularité d'espacement des points. En effet, le contour doit conserver une forme arrondie en minimisant les dérivées d'ordre 1, 2, ... et doit empêcher un point de se détacher trop loin du reste du contour. Idéalement, l'énergie interne est minimale pour un cercle où tous les points sont régulièrement espacés.

Grande liberté est laissée au concepteur pour imaginer toutes les heuristiques qu'il voudra [PS08].

Appelée aussi énergie de régularisation, elle permet un contrôle de la forme de notre courbe. Elle se compose de deux termes :

Terme 1 : il est fonction de la première dérivée de la courbe et agit essentiellement sur la longueur de cette dernière. Sa formulation mathématique est la suivante :

$$\left| \frac{\partial v(s, t)}{\partial x} \right|^2 \quad (I.8)$$

Terme 2 : il est fonction de la seconde dérivée de la courbe et agit essentiellement sur la courbure. Sa formulation mathématique est la suivante :

$$\left| \frac{\partial^2 v(s, t)}{\partial x^2} \right|^2 \quad (I.9)$$

Les deux termes sont multipliés par deux coefficients de pondération, respectivement de lissage ($\alpha \geq 0$) et de rigidité ($\beta \geq 0$). Lorsque $\alpha = 0$, la courbe peut présenter des discontinuités. On parlera donc d'énergie de continuité. Lorsque $\beta = 0$, la courbe peut prendre une forte convexité, par contre lorsque β est grand, la courbe tendra vers un cercle ou un rectangle si elle fermée ou une droite si elle est ouverte [BOU09].

$$E = \alpha \int_a^b \left| \frac{\partial v(s, t)}{\partial x} \right|^2 ds + \beta \left| \frac{\partial v^2(s, t)}{\partial x^2} \right|^2 ds \quad (I.10)$$

C.Énergie De Contexte

L'énergie de contexte, parfois appelée de contrainte, permet d'introduire des connaissances a priori sur ce que nous cherchons. Entre autres, nous plaçons, sous cette rubrique, l'énergie ballon introduite par Laurent D. Cohen.

La minimisation de l'énergie implique une minimisation de distance. La force ballon va tendre à gonfler le contour actif ou accélérer sa rétraction selon le signe de la force introduite. De plus, cette force va permettre de dépasser les contours présentant un faible gradient et ainsi de sortir du bruit pour atteindre une frontière plus fortement marquée. Il s'agit d'une force normale au contour en chaque point.

$$F_{ballon}(v(s)) = k\vec{n}(s) \quad (I.11)$$

Où $\vec{n}(s)$ est un vecteur unitaire normal à la courbe en $v(s)$. L'intensité de l'énergie ballon est un scalaire généralement négatif (expansion du ballon) proportionnel à l'aire intérieur du contour.

C'est à partir de ces différentes énergies et de leur combinaison que l'on peut définir un problème dont la solution est le contour recherché. La mise en œuvre d'une telle approche a donné lieu de nombreuses implémentations [BOU09].

I.8.3.6 Algorithme "Greedy"

L'utilisation de l'algorithme Greedy pour minimiser l'énergie d'un contour actif, comme l'ont proposé Williams et Shah est devenue une alternative assez fréquente à l'approche variationnelle. Williams et Shah, de manière classique et comme Kass et Amini, Discrétisent l'expression qui Calcul l'énergie totale par algorithme de Greedy :

$$E_{tot} = \int_0^1 E_{ext}(v(s)) + \frac{1}{2}(\alpha(s)|v'(s)|^2 + \beta(s)|v''(s)|^2) ds \quad (I.12)$$

Par différences finies et utilisant pour la continuité :

$$\left\| \frac{dv_i}{ds} \right\|^2 = \|v_i - v_{i-1}\|^2 = (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 \quad (\text{I.13})$$

Et pour la courbure :

$$\left\| \frac{d^2v_i}{ds^2} \right\|^2 = \|v_{i+1} - 2v_i + v_{i-1}\|^2 = (x_{i-1} - 2x_i + x_{i+1})^2 + (y_{i-1} - 2y_i + y_{i+1})^2 \quad (\text{I.14})$$

Il est à remarquer que ces deux expressions supposent deux hypothèses. La première est que les points soient placés, le long de la courbe, à une distance unité les uns des autres. Si les points sont équi-répartis, alors la première équation doit être divisée par d^2 et la deuxième équation par d^4 dans le cas contraire, continuité est incorrecte d'un facteur d_i^2 où d_i correspond à la distance entre les points i et $i - 1$.

Cela impliquera une valeur supérieure de la continuité, de même pour la courbure. La seconde hypothèse suppose que le paramètre soit une longueur d'arc, pour que la courbure soit donnée par $\|v_{ss}\|$.

L'algorithme du Greedy, comme la programmation dynamique autorise l'introduction d'autres contraintes. Il est plus rapide que la méthode précédente puisqu'il est, à chaque itération en $O(nm)$ Au lieu de $O(nm^3)$, pour un contour comportant n points et un voisinage de taille m .

L'algorithme est itératif comme celui de Kass et Amini. On examine, à chaque itération, le voisinage de chaque point. On choisit le point du voisinage qui donne la plus faible valeur de l'énergie totale. Dans le cas d'une approche discrète, le contour est modélisé par un ensemble de points ordonnés $M_i, i \in [1, N]$ l'énergie de contour alors assimilé à la somme des énergies associées au N . Définissant le contour. (L'énergie totale par algorithme de Greedy dans le cas Discret)[ROU03].

$$E(e) = \sum_{i=1}^N (\alpha E_{\text{continuité}}(M_i) + \beta E_{\text{courbure}}(M_i) + \gamma E_{\text{gradient}} + \delta E_{\text{ballon}}(M_i)) \quad (\text{I.15})$$

I.9 Conclusion

Dans ce chapitre, quelques notions fondamentales liées au traitement d'image et la vision artificielle ont été abordés ainsi que les différentes étapes de traitement et prétraitement d'image, nous avons fait un tour assez exhaustif des méthodes de détection de contours dans une image 2D où de nombreux progrès restent à faire car il n'existe pas encore d'algorithme assez robuste permettant de détecter n'importe quel type de contour dans une image. Rappelons qu'un contour peut être considéré comme une frontière d'une région. Nous avons vite constaté la faiblesse de ces méthodes due au bruit contenu dans une image et aux contours parasites. Une amélioration possible de ces méthodes est de tenir compte des particularités locales en utilisant un seuillage par Hystérésis qui permet ainsi de repérer les maxima globaux et d'éviter les maxima locaux. Nous avons étudié aussi le principe des contours actifs (ou snakes) en présentant son état de l'art.

Dans les deux prochains chapitres, on va décrire de manière détaillée une des techniques de détection de contours à savoir les contours actifs génétique.

Chapitre 02 : Les Algorithmes Génétiques

II.1 Introduction

Les algorithmes génétiques sont des algorithmes d'exploration fondés sur les mécanismes de la sélection naturelle et de la génétique. Ils utilisent à la fois les principes de la survie des structures les mieux adaptées, et les échanges d'information pseudo-aléatoires, pour former un algorithme d'exploration qui possède certaines des caractéristiques de l'exploration humaine. Ils sont une transcription directe du processus naturel d'évolution au monde de l'informatique. Les algorithmes génétiques sont des algorithmes d'optimisation s'appuyant sur des techniques dérivées de la génétique et des mécanismes d'évolution de la nature : croisements, mutations, sélections, etc... Ils appartiennent à la classe des algorithmes évolutionnaires.

II.2 Historique

C'est en 1860 que Charles Darwin publie son livre intitulé *L'origine des espèces au moyen de la sélection naturelle ou la lutte pour l'existence dans la nature*. Dans ce livre, Darwin rejette l'existence «de systèmes naturels figés», déjà adaptés pour toujours à toutes les conditions extérieures, et expose sa théorie de l'évolution des espèces : sous l'influence des contraintes extérieures, les êtres vivants se sont graduellement adaptés à leur milieu naturel au travers de processus de reproductions.

Darwin proposa une théorie qui clarifie l'évolution des espèces en mettant en avant quatre lois :

- La loi de croissance et de reproduction.
- La loi d'hérédité qu'implique quasiment la loi de reproduction
- La loi de variabilité, résultant des conditions d'existence.
- La loi de multiplication des espèces qui amène la lutte pour l'existence et qui a pour conséquence la sélection naturelle.

Presque simultanément, en 1866, Mendel (« le moine des pois ») publie l'article retraçant dix années d'expériences d'hybridation chez les végétaux (recombinaison des gènes) et l'adresse aux sociétés scientifiques des quatre coins du monde. Les réactions sont mitigées, voire inexistantes.

Le monde scientifique n'est pas prêt à reconnaître la qualité de ses résultats. Ce n'est seulement en 1900, que la publication de trois nouveaux articles signés Hugo de Vries, Carl Correns et Erich von Tschermak révèle des résultats similaires à ceux de Mendel, et feront que ces premiers seront reconnus.

Au début des années 1960 que John Holland de l'Université du Michigan a commencé à s'intéresser à ce qui allait devenir les algorithmes génétiques. Ces algorithmes font partie du champ de l'Intelligence Artificielle (IA). Il s'agit d'IA de bas niveau, inspirée par « l'intelligence » de la nature. Ils sont basés sur le concept de sélection naturelle élaboré par Charles Darwin. Le vocabulaire employé est donc directement calqué sur celui de la théorie de l'évolution et de la génétique, John Holland étudie les systèmes évolutifs et, en 1975, il introduit le premier modèle formel des algorithmes génétiques (the canonical génétique algorithme AGC) dans son livre *Adaptation in Natural and Artificial Systems*. Il expliqua comment ajouter de l'intelligence dans un programme informatique avec les croisements (échangeant le matériel génétique) et la mutation (source de la diversité génétique)[SR04].

II.3 Analogie Avec La Biologie

Les algorithmes génétiques étant basés sur des phénomènes biologiques, il convient de rappeler au préalable quelques termes de génétique [3].

II.3.1 Les Bases des Algorithmes Génétique

Les organismes vivants sont constitués de cellules, dont les noyaux comportent des chromosomes qui sont des chaînes d'ADN. L'élément de base de ces chromosomes (le caractère de la chaîne d'ADN) est un gène (figure (2.1)). Sur chacun de ces chromosomes, une suite de gènes constitue une chaîne qui code les fonctionnalités de l'organisme (la couleur des yeux...etc.). La position d'un gène sur le chromosome est son locus. L'ensemble des gènes d'un individu est son génotype et l'ensemble du patrimoine génétique d'une espèce est le génome. Les différentes versions d'un même gène sont appelées allèles.

On utilise aussi, dans les algorithmes génétiques, une analogie avec la biologie, qui concerne l'évolution, hypothèse émise par Darwin et qui propose qu'au fil du temps, les gènes conservés au sein d'une population donnée sont ceux qui sont le plus adaptés aux besoins de l'espèce vis à vis de son environnement [3].

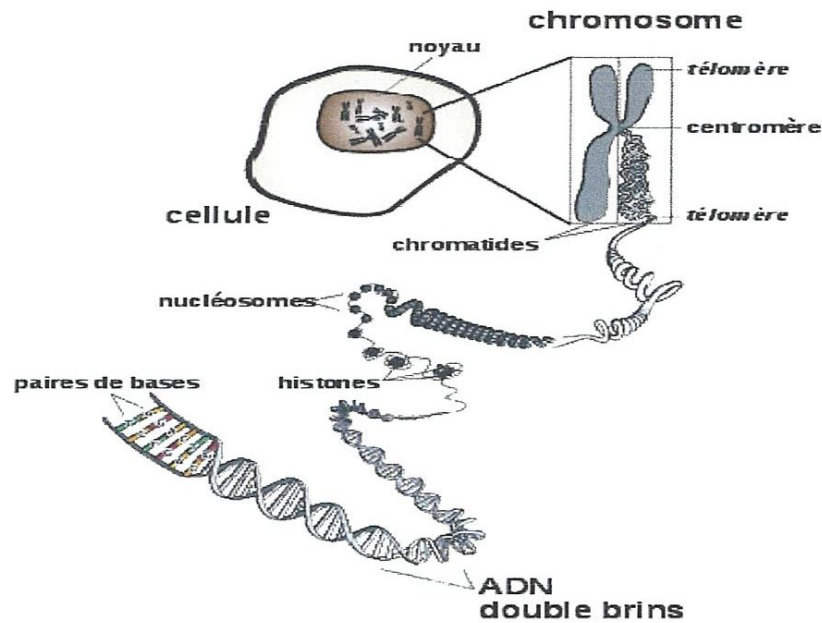


Figure 2.1 : Chromosome humain.

On utilise aussi, dans les algorithmes génétiques, une analogie avec la biologie, qui concerne l'évolution, hypothèse émise par Darwin et qui propose qu'au fil du temps, les gènes conservés au sein d'une population donnée sont ceux qui sont le plus adaptés aux besoins de l'espèce vis à vis de son environnement [3].

II.3.2 Les Outils Issus De La Biologie :

La génétique a mis en évidence l'existence de plusieurs opérations au sein d'un organisme donnant lieu au brassage génétique. Ces opérations interviennent lors de la phase de reproduction lorsque les chromosomes de deux organismes fusionnent.

Ces opérations sont imitées par les algorithmes génétiques afin de faire évoluer les populations de solutions de manière progressive [3].

II.3.2.1 La sélection

Pour déterminer quels individus sont plus enclins à obtenir les meilleurs résultats, une sélection est opérée. Ce processus est analogue à un processus de sélection naturelle, les individus les plus adaptés gagnent la compétition de la reproduction tandis que les moins adaptés meurent avant la reproduction, ce qui améliore globalement l'adaptation. Etant donné que la sélection est le résultat d'une intervention humaine ou, du moins, l'application d'un critère défini par l'homme, les algorithmes génétiques devraient donc plutôt être rapprochés

de la sélection artificielle telle que la pratiquent les agriculteurs que de la sélection naturelle, qui œuvre "en aveugle".

II.3.2.2 Les Croisement Ou Recombinaison

Lors de cette opération, deux chromosomes s'échangent des parties de leurs chaînes, pour donner de nouveaux chromosomes. Ces croisement peuvent être simples ou multiples.

Dans le premier cas, les deux chromosomes se croisent et s'échangent des portions d'ADN en un seul point. Dans le deuxième cas, il y a plusieurs points de croisement. Pour les algorithmes génétiques, c'est cette opération (le plus souvent sous sa forme simple) qui est prépondérante. Sa probabilité d'apparition lors d'un croisement entre deux chromosomes est un paramètre de l'algorithme génétique et dépend du problème et de la technique de recombinaison. La probabilité d'un enjambement est alors comprise entre 0 et 1 strictement [3].

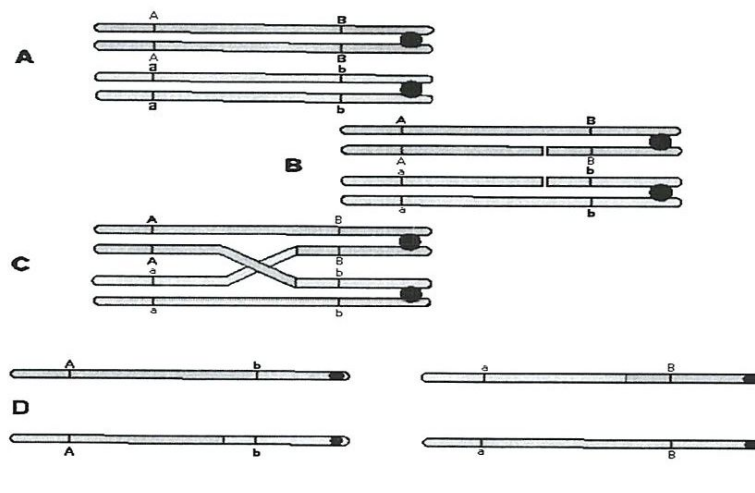


Figure 2.2 : Croisement humain

II.3.2.3 Les mutations

De façon aléatoire, un gène peut, au sein d'un chromosome être substitué à un autre. De la même manière que pour les enjambements, on définit ici un taux de mutation lors des changements de population qui est généralement compris entre 0,001 et 0,01. Il est nécessaire de choisir pour ce taux une valeur relativement faible de manière à ne pas tomber dans une recherche aléatoire et conserver le principe de sélection et d'évolution. La mutation sert à éviter une convergence prématurée de l'algorithme. Par exemple lors d'une recherche d'extremum la mutation sert à éviter la convergence vers un extremum local[3].

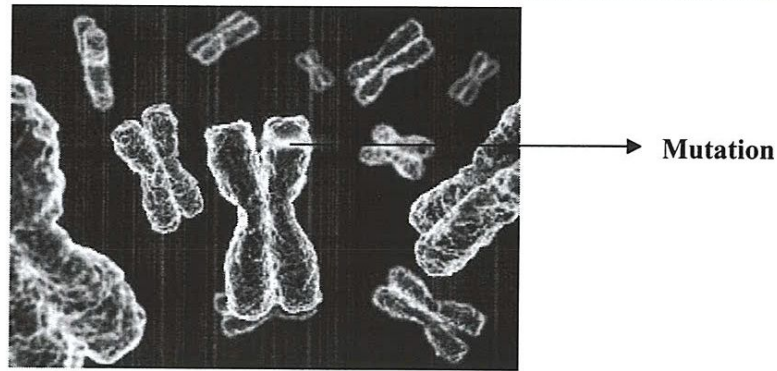


Figure 2.3 : Mutation humaine.

II.4 Le Principe des Algorithmes Génétiques

A chaque génération, l'AG standard décrit originellement par J.Holland (Goldberg, 1989 ; Holland, 1992 ; Man et al., 1996 ; Schmitt, 2001 ; Petrowski) accomplit une génération lorsque les opérateurs génétiques ont été appliqués simultanément aux individus de la population, en effet, à chaque génération, un nouvel ensemble de créatures artificielles (des chaînes de caractères) est créé en utilisant des parties des meilleurs éléments de la génération précédente. Bien qu'utilisant le hasard, les AGs ne sont pas purement aléatoires. Ils exploitent efficacement l'information obtenue précédemment pour spéculer sur la position de nouveaux points à explorer, avec l'espoir d'améliorer la performance. Nous donnons les différentes étapes d'un AGS dans la figure suivante [LA91].

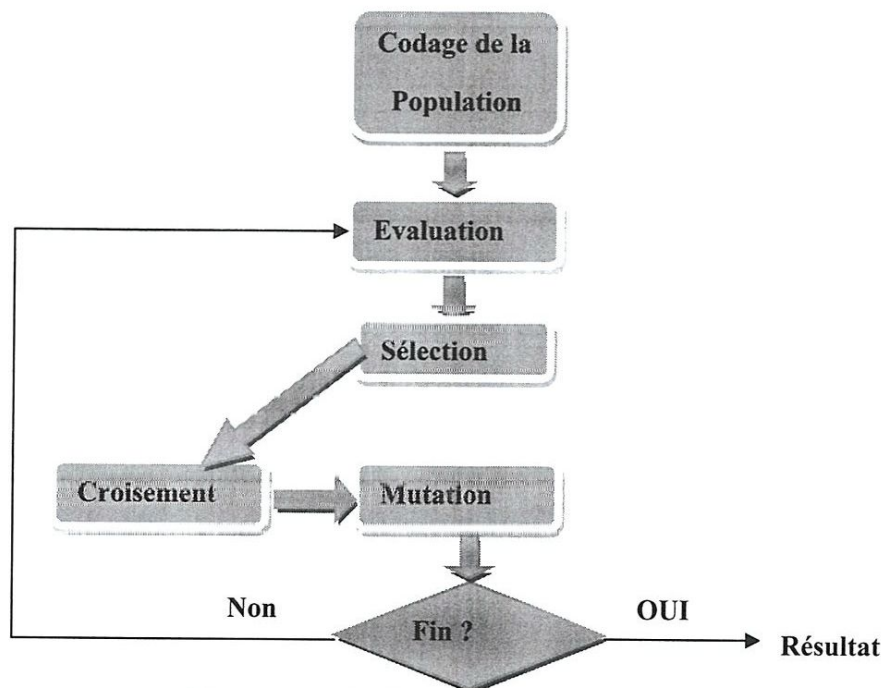


Figure 2.4 : les étapes d'un algorithme génétique.

II.4.1 Le Codage De La Population

Chaque paramètre d'une solution est assimilé à un gène, toutes les valeurs qu'il peut prendre sont les allèles de ce gène, on doit trouver une manière de coder chaque allèle différent de façon unique (établir une bijection entre l'allèle "réel" et sa représentation codée).

Un chromosome est une suite de gène, on peut par exemple choisir de regrouper les paramètres similaires dans un même chromosome (chromosome à un seul brin) et chaque gène sera repérable par sa position : son locus sur le chromosome en question.

Chaque individu est représenté par un ensemble de chromosomes, et une population est un ensemble d'individus[SR04].

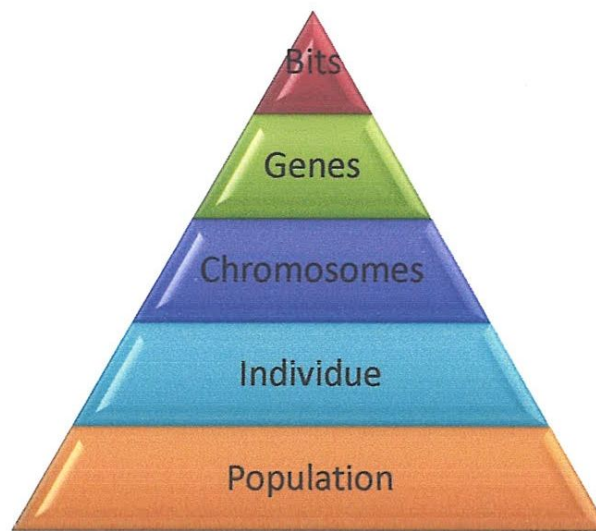


Figure 2.5 : les niveaux d'organisation d'un algorithme génétique.

II.4.1.1 Les Types De Codage

C'est un moyen de coder les solutions d'un problème ou encore les individus. Chaque solution est constituée d'une chaîne de symboles appartenant à un alphabet à priori de petite taille. Le symbole a un rôle dans les AGs similaire au rôle du gène dans la génétique naturelle. Le choix d'un code doit tenir compte de la complexité du processus de codage/décodage qui peut ralentir les calculs et influencer considérablement la convergence de l'AG. Il y a trois principaux types de codage utilisables, et on peut passer de l'un à l'autre relativement facilement :

Le Codage Binaire

C'est le plus utilisé, Chaque gène dispose du même alphabet binaire $\{0, 1\}$ Un gène est alors représenté par un entier long (32 bits), les chromosomes qui sont des suites de gènes sont représentés par des tableaux de gènes et les individus de notre espace de recherche sont représentés par des tableaux de chromosomes. Ce cas peut être généralisé à tout alphabet allélique n-aire permettant un codage plus intuitif Par exemple pour le problème du voyageur de commerce on peut préférer utiliser l'alphabet Allélique $\{c_1, c_2, c_3, \dots, c_n\}$ où c_i représente la ville de numéro i [SR04].

Le Codage Réel

Cela peut-être utile notamment dans le cas où l'on recherche le maximum d'une fonction réelle comme c'est présenté dans la figure ci-dessous :

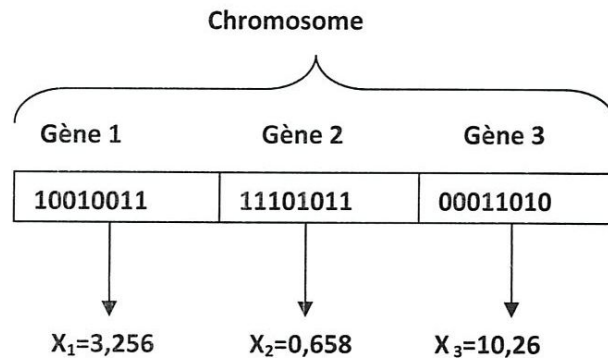


Figure 2.6 : Illustration schématique du codage des variables réelle.

Le Codage de Gray

Dans le cas d'un codage binaire on utilise souvent la "distance de Hamming" comme mesure de la dissimilarité entre deux éléments de population, cette mesure compte les différences de bits de même rang de ces deux séquences. Et c'est là que le codage binaire commence à montrer ses limites. En effet, deux éléments voisins en termes de distance de Hamming ne codent pas nécessairement deux éléments proches dans l'espace de recherche. Cet inconvénient peut être évité en utilisant un "codage de Gray" : le codage de Gray est un codage qui a comme propriété qu'entre un élément n et un élément $n+1$, donc voisin dans l'espace de recherche, un seul bit diffère [SR04].

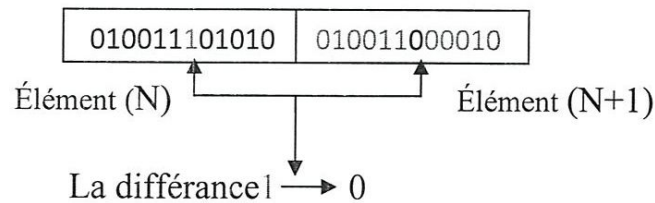


Figure 2.7 : le codage en gray.

II.4.2 La Fonction D'évaluation

Elle prend en argument l'individu et lui associe un coût de performance. Cette fonction doit guider l'algorithme vers l'optimum en réalisant implicitement une pression de sélection dans cette direction. De plus, l'efficacité d'un AG s'appuie pour une grande part, sur la qualité de cette fonction. La mise au point d'une bonne fonction d'adaptation doit respecter plusieurs critères qui se rapportent à sa complexité et à la satisfaction des contraintes du problème. Lorsque la fonction d'adaptation apparaît excessivement complexe, consommant une importante puissance de calcul, la recherche d'une approximation plus simple est souhaitable. L'évaluation de la qualité (fitness) d'un individu permet d'illustrer avec une valeur numérique (et donc plus facile à manipuler) la qualité des gènes qui forment l'individu. Plus la qualité d'un individu est grande, plus ses chances d'être sélectionné pour la reproduction sont grandes. En général, on calcule les chances de reproduction d'un individu en regard de sa qualité en relation avec la qualité totale des individus de la population [LA91].

Dans cette figure on a représenté la fonction d'évaluation la fonction qui donne une valeur pour chaque individu par rapport a son degré de d'adaptation (Figure 2.8).

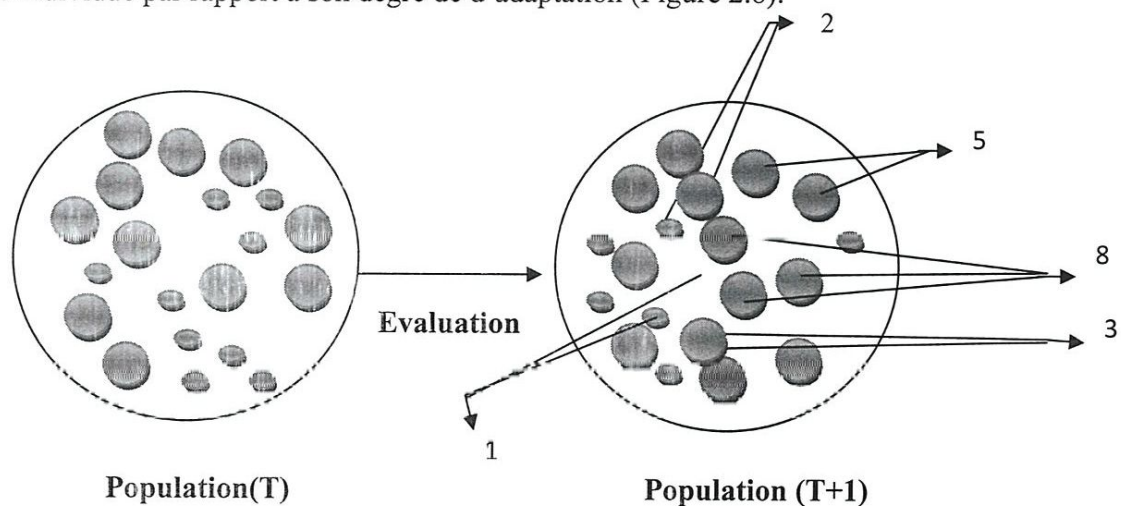


Figure 2.8 :La Fonction D'évaluation.

II.4.3 La Sélection :

Cette opération est fondée sur le principe d'adaptation de chaque individu d'une Population à son environnement, suivant la théorie de la sélection naturelle introduite Par Charles Darwin. Ainsi, seuls les individus les plus aptes à répondre à certains critères seront sélectionnés à survivre et à se multiplier. D'autres individus prendront la place des plus mauvais.

La fonction d'évaluation de l'adéquation des individus évalue chaque chromosome de la population, et seuls ceux dont la qualité sera suffisante seront conservés d'une génération à une autre [LA91].

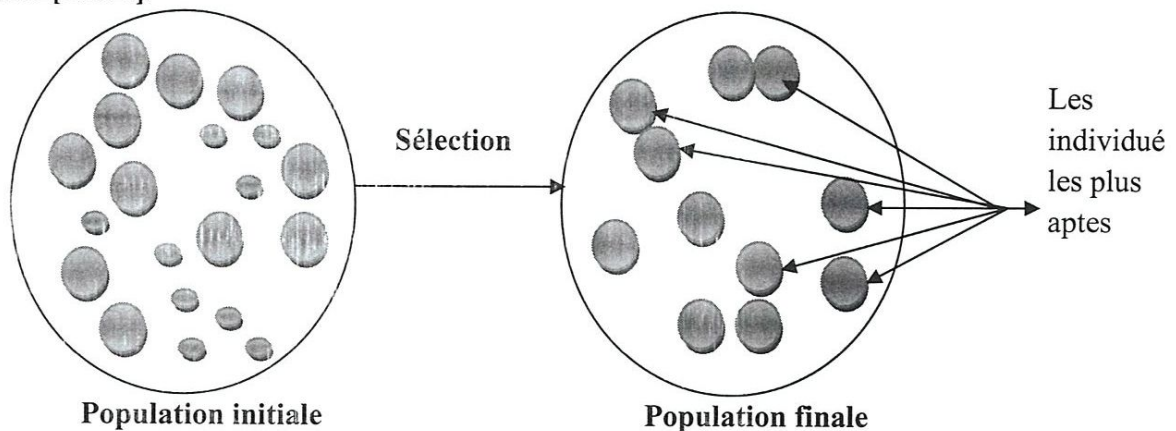


Figure 2.9 : La Sélection.

II.4.3.1 L'opérateur de sélection

Cet opérateur est chargé de définir quels seront les individus de P qui vont être dupliqués dans la nouvelle population P' et vont servir de parents (application de l'opérateur de croisement).

Soit n le nombre d'individus de P , on doit en sélectionner $n/2$ (l'opérateur de croisement nous permet de repasser à n individus).

Cet opérateur est peut-être le plus important puisqu'il permet aux individus d'une population de survivre, de se reproduire ou de mourir.

En règle générale, la probabilité de survie d'un individu sera directement liée à son efficacité relative au sein de la population [SR04].

II.4.3.2 Les méthodes de sélection :

On trouve essentiellement quatre types de méthodes de sélection différentes. Le choix de la bonne méthode dépend du contexte d'utilisation :

- La méthode de la "loterie biaisée" (roulette wheel) de Goldberg,
- La méthode "élitiste",
- La sélection par tournois,
- La sélection universelle stochastique [SR04].

A- La Loterie Biaisée

C'est comme au loto (hasard), Cette méthode est la plus connue et la plus utilisée, avec cette méthode chaque individu a une chance d'être sélectionné proportionnelle à sa performance, donc plus les individus sont adaptés au problème, plus ils ont de chances d'être sélectionnés.

Pour utiliser l'image de la "roue du forain", chaque individu se voit attribué un secteur dont l'angle est proportionnel à son adaptation, sa "fitness".

On fait tourner la roue et quand elle cesse de tourner on sélectionne l'individu correspondant au secteur désigné par une sorte de "curseur", curseur qui pointe sur un secteur particulier de celle-ci après qu'elle se soit arrêtée de tourner [SR04].

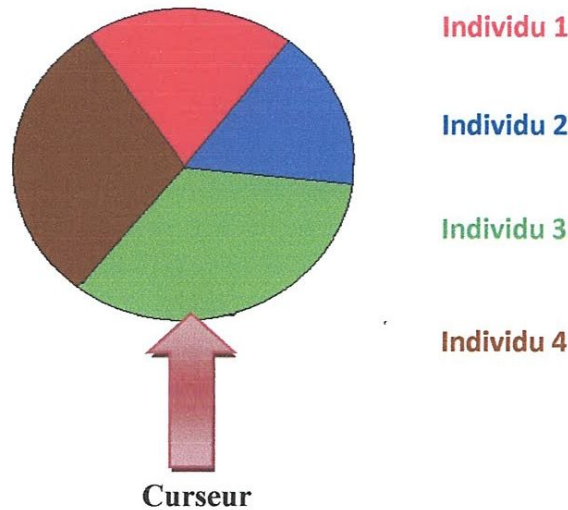


Figure 2.10 : la méthode de sélection de la loterie biaisée.

Cette méthode, bien que largement répandue, a pas mal d'inconvénients :

En effet, elle a une forte variance. Il n'est pas impossible que sur n sélections successives destinées à désigner les parents de la nouvelle génération P' , la quasi-totalité, voire pire la totalité des n individus sélectionnés soient des individus ayant une fitness vraiment mauvaise et donc que pratiquement aucun individu voire aucun individu à forte fitness ne fasse partie des parents de la nouvelle génération. Ce phénomène est bien sûr très dommageable car cela va complètement à l'encontre du principe des algorithmes génétiques qui veut que les meilleurs individus soient sélectionnés de manière à converger vers une solution la plus optimale possible.

A l'inverse, on peut arriver à une domination écrasante d'un individu "localement supérieur".

Ceci entraînant une grave perte de diversité. Imaginons par exemple qu'on ait un individu ayant une fitness très élevée par rapport au reste de la population, disons dix fois supérieure, il n'est pas impossible qu'après quelques générations successives on se retrouve avec une population ne contenant que des copies de cet individu. Le problème est que cet individu avait une fitness très élevée, mais que cette fitness était toute relative, elle était très élevée mais seulement en comparaison des autres individus. On se retrouve donc face à problème connu sous le nom de "convergence prématurée"; l'évolution se met donc à stagner et on atteindra alors jamais l'optimum, on restera bloqué sur un optimum local.

Il existe certaines techniques pour essayer de limiter ce phénomène, comme par exemple le "scaling", qui consiste à effectuer un changement d'échelle de manière à augmenter ou diminuer de manière forcée la fitness d'un individu par rapport à un autre selon leur écart de fitness. Malgré tout, il est conseillé d'opter plutôt pour une autre méthode de sélection[SR04].

B- La Méthode Elitiste

Cette méthode consiste à sélectionner les n individus dont on a besoin pour la nouvelle génération P' en prenant les n meilleurs individus de la population P après l'avoir triée de manière décroissante selon la fitness de ses individus.

Il est inutile de préciser que cette méthode est encore pire que celle de la loterie biaisée dans le sens où elle amènera à une convergence prématurée encore plus rapidement et surtout de manière encore plus sûre que la méthode de sélection de la loterie biaisée ; en effet, la pression de la sélection est trop forte, la variance nulle et la diversité inexistante, du moins le peu de diversité qu'il pourrait y avoir ne résultera pas de la sélection mais plutôt du croisement et des mutations. Là aussi il faut opter pour une autre méthode de sélection[SR04].

C- La Sélection Par K-Tournois

Cette méthode est celle avec laquelle on obtient les résultats les plus satisfaisants. Le principe de cette méthode est le suivant : on effectue un tirage avec remise de deux individus de P , et on les fait "combattre". Celui qui a la fitness la plus élevée l'emporte avec une probabilité p comprise entre 0.5 et 1. On répète ce processus n fois de manière à obtenir les n individus de P' qui serviront de parents[SR04].

La variance de cette méthode est élevée et le fait d'augmenter ou de diminuer la valeur de p permet respectivement de diminuer ou d'augmenter la pression de la sélection.

D- La Sélection Universelle Stochastique

Cette méthode semble être très peu utilisée et qui plus est possède une variance faible, donc introduit peu de diversité, nous n'entrerons donc pas dans les détails, on se contentera

d'exposer sa mise en œuvre : On prend l'image d'un segment découpé en autant de sous-segments qu'il y a d'individus [SR04].

Les individus sélectionnés sont désignés par un ensemble de points équidistants.

II.4.4 Le Croisement (Cross Over)

Le croisement a pour but d'enrichir la diversité de la population en manipulant la structure des chromosomes. Classiquement, les croisements sont envisagés avec deux parents et génèrent deux enfants.

Les croisements permettent de simuler des reproductions d'individus dans le but d'en créer de nouveaux.

Initialement, le croisement associé au codage par chaînes de bits est le croisement à découpage de chromosomes (slicing crossover). Pour effectuer ce type de croisement sur des chromosomes constitués de M gènes, on tire aléatoirement une position dans chacun des parents. On échange ensuite les deux sous-chaînes terminales de chacun des deux chromosomes, ce qui produit deux enfants $C1$ et $C2$ (voir la Figure (2.11)) [JN05].

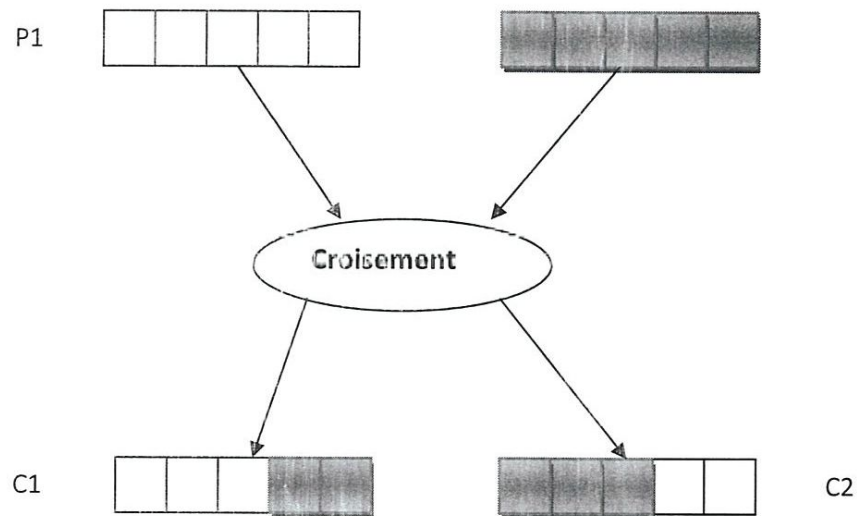


Figure 2.11 : Le Croisement (Cross Over).

II.4.4.1 L'opérateur de croisement

Le crossover utilisé par les algorithmes génétiques est la transposition informatique du mécanisme qui permet, dans la nature, la production de chromosomes qui héritent partiellement des caractéristiques des parents.

Son rôle fondamental est de permettre la recombinaison des informations présentes dans le patrimoine génétique de la population.

Cet opérateur est appliqué après avoir appliqué l'opérateur de sélection sur la population P ; on se retrouve donc avec une population P' de $n/2$ individus et on doit doubler ce nombre pour que notre nouvelle génération soit complète. (ou n nombre d'individu).

On va donc créer de manière aléatoire $n/4$ couples et on les fait se "reproduire".

Les chromosomes (ensembles de paramètres) des parents sont alors copiés et recombinaison de façon à former deux descendants possédant des caractéristiques issues des deux parents.

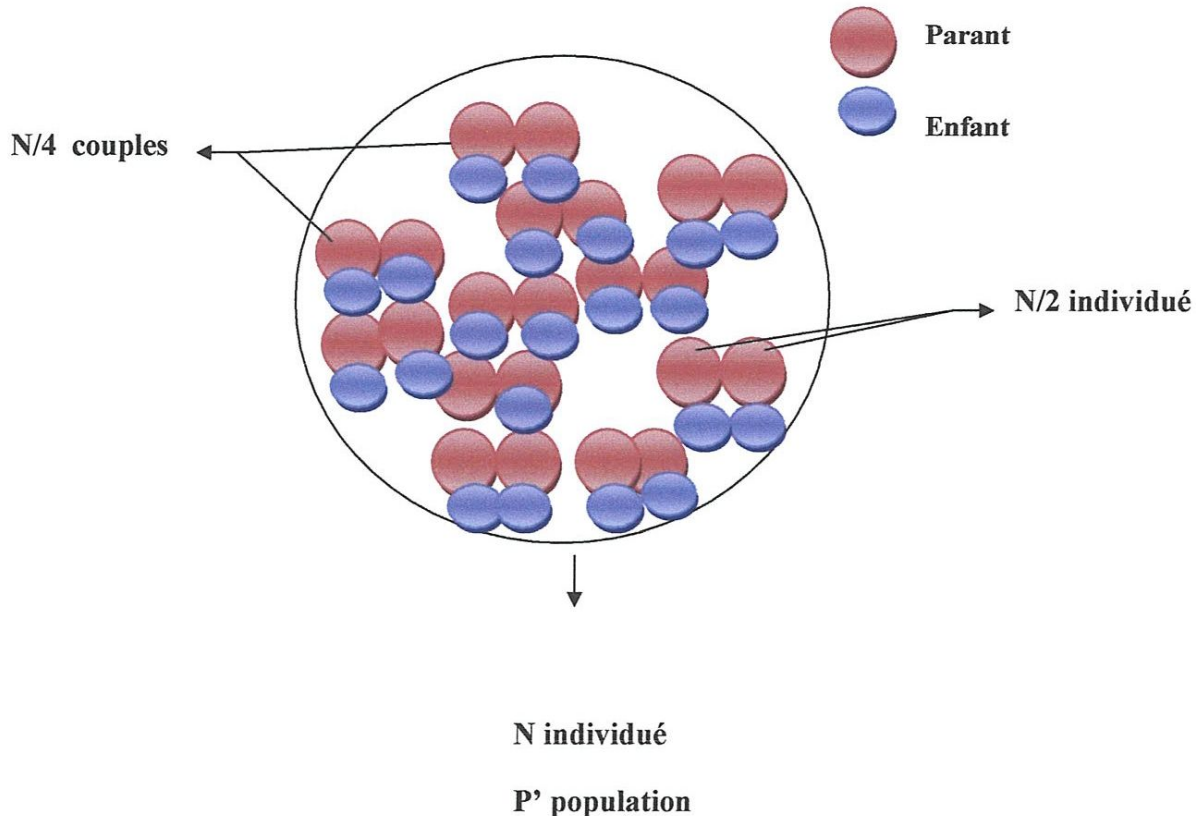


Figure 2.12 : L'opérateur de croisement.

Dans chaque couple au niveau de leur chromosome en fixe au hasard un ou plusieurs points de croisement (loci) chaque chromosome se retrouve donc séparé en "segments". Puis chaque segment du parent 1 est échangé avec son "homologue" du parent 2 selon une probabilité de croisement p_c . De ce processus résulte 2 fils pour chaque couple et notre population P' contient donc bien maintenant n individus (Figure (2.12)).

On peut noter que le nombre de points de croisements ainsi que la probabilité de croisement p_c permettent d'introduire plus ou moins de diversité.

En effet, plus le nombre de points de croisements sera grand et plus la probabilité de croisement sera élevée plus il y aura d'échange de segments, donc d'échange de paramètres,

d'information, et plus le nombre de points de croisements sera petit et plus la probabilité de croisement sera faible, moins le croisement apportera de diversité[SR04].

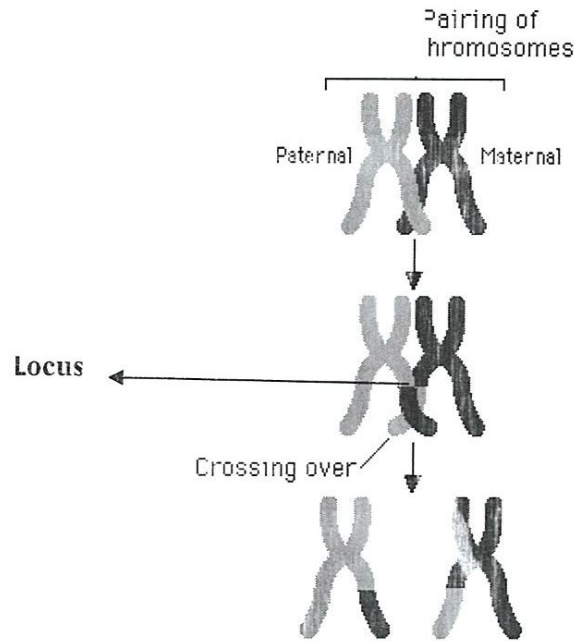


Figure 2.13 : le croisement des chromosomes.

II.4.4.2 Les Types De Croisement

Le croisement consiste à combiner deux individus quelconques (dits parents) pour en ressortir deux autres individus (dits enfants) pas forcément meilleurs que les parents ; il existe plusieurs variantes de cet opérateur, mais en général il consiste à couper en un ou plusieurs points deux individus (aux mêmes endroits dans les deux individus) et à échanger les parties situées entre ces points, ces derniers sont générés aléatoirement [LA91].

A- Croisement En Un Point :

Sa consiste a coupé les chromosomes des individu (les parlants) en un seule point choisi au hasard dans la même zone pour les 2 chromosomes, après en procède a l'échange des segments entre eux pour donné la naissance des deux enfants(Figure (2.14)).

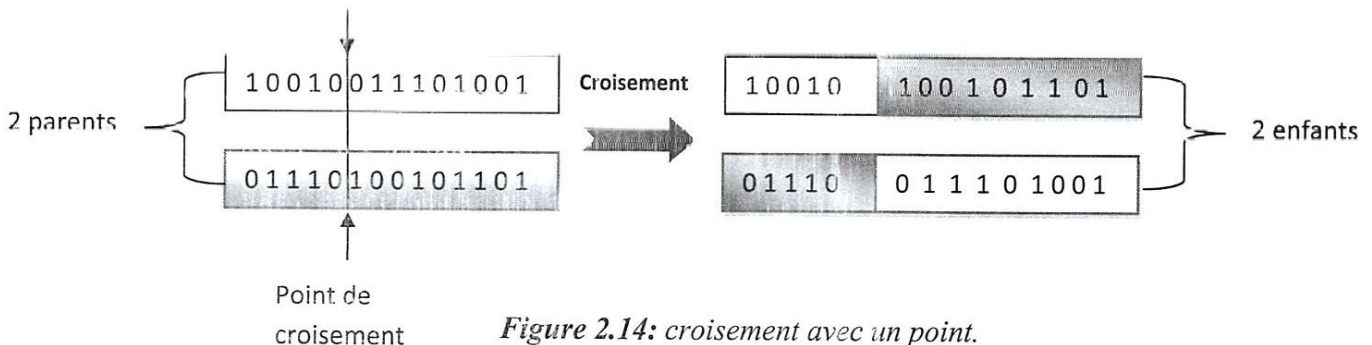


Figure 2.14: croisement avec un point.

B- Le Croisement Multi Point

En par de même principe de la premier approche (1-point) mais dans cette fois en coupe dans deux point identique le chromosome des parents, avec deux point ou pulse de croisement en peut avoir trois segment de chromosome qui échange entre eux(Figure (2.15)).

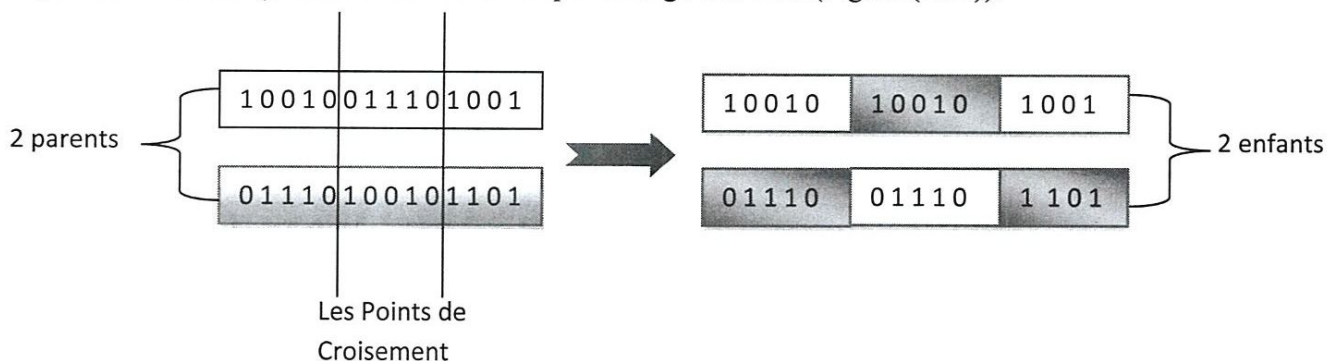


Figure 2.15: croisement multi point.

C- Le Croisement Uniforme

Une autre variante est le croisement uniforme qui utilise une chaîne de bits générée aléatoirement et de même longueur que les individus appelée masque. Les gènes des individus initiaux sont échangés en fonction de cette chaîne aléatoire, lorsque le bit correspondant =1. Le croisement uniforme de permutations a les avantages de la simplicité et selon son auteur (Syswerda, 1989), d'une bonne efficacité[LA91].

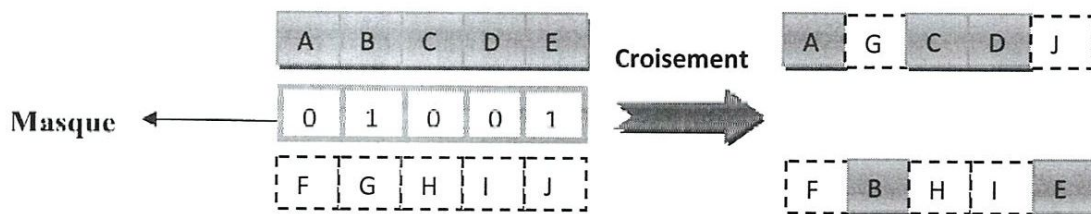


Figure 2.16: croisement uniforme.

II.4.5La Mutation

Une autre solution que le croisement pour créer de nouveaux individus est de modifier ceux déjà existants. Une fois de plus, le hasard va nous être d'une grande utilité. Il peut s'avérer efficace de modifier aléatoirement quelques individus de notre population en modifiant un gène ou un autre. Rien ne nous dit que l'individu mute sera meilleur ou moins bon, mais il apportera des possibilités supplémentaires qui pourraient bien être utiles pour la création de bonnes solutions. De même que pour les croisements, il n'est pas recommandé de faire muter tous les individus. Il est possible de faire muter un individu de la manière qu'il vous plaira. Une seule contrainte : l'individu muté doit être de la forme d'une solution potentielle [4].

II.4.5.1 L'opérateur De Mutation

Cet opérateur agit sur un individu. Il consiste à choisir d'une manière aléatoire un ou plusieurs gènes (bits) et à modifier leurs valeurs. Chaque bit d'un individu a une probabilité pour qu'il subisse une mutation, notée p_m .

Une mutation consiste simplement en l'inversion d'un bit (ou de plusieurs bits, mais vu la probabilité de mutation c'est extrêmement rare) se trouvant en un locus bien particulier et lui aussi déterminé de manière aléatoire; on peut donc résumer la mutation de la façon suivante :

On utilise une fonction censée nous retourner *true* avec une probabilité p_m .

Pour chaque locus **faire**
 Faire appel à la fonction
Si cette fonction nous renvoie *true* **alors**
 on inverse le bit se trouvant à ce locus
FinSi
FinPour

Plusieurs arguments justifient l'utilisation de la mutation. Ainsi, les actions combinées de la pression de sélection et du croisement peuvent faire disparaître des gènes, aux mêmes positions des génotypes pour tous les individus. La mutation est capable de réintroduire ces gènes dans la population et participe ainsi au maintien de la diversité.

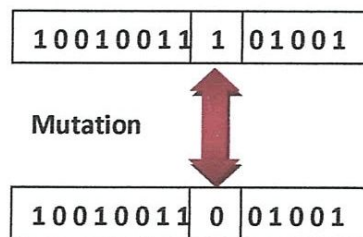


Figure 2.17 : L'opérateur De Mutation.

L'opérateur de mutation modifie donc de manière complètement aléatoire les caractéristiques d'une solution, ce qui permet d'introduire et de maintenir la diversité au sein de notre population de solutions. Cet opérateur joue le rôle d'un "élément perturbateur", il introduit du "bruit" au sein de la population.

Cet opérateur dispose de quatre grands avantages :

- Il garantit la diversité de la population, ce qui est primordial pour les algorithmes génétiques.

- Il permet d'éviter un phénomène connu sous le nom de *dérive génétique*. On parle de dérive génétique quand certains gènes favorisés par le hasard se répandent au détriment des autres et sont ainsi présents au même endroit sur tous les chromosomes. Le fait que l'opérateur de mutation puisse entraîner de manière aléatoire des changements au niveau de n'importe quel locus permet d'éviter l'installation de cette situation défavorable.
- Il permet de limiter les risques d'une convergence prématurée causée par exemple par une méthode de sélection élitiste imposant à la population une pression sélective trop forte. En effet, dans le cas d'une convergence prématurée on se retrouve avec une population dont tous les individus sont identiques mais ne sont que des optimums locaux. Tous les individus étant identiques, le croisement ne changera rien à la situation. En effet, l'échange d'informations par crossover entre des individus strictement identiques est bien sûr totalement sans conséquences; on aura beau choisir la méthode de croisement qu'on veut on se retrouvera toujours à échanger des portions de chromosomes identiques et la population n'évoluera pas. L'évolution se retrouvant bloquée on n'attendra jamais l'optimum global.

La mutation entraînant des inversions de bits de manière aléatoire permet de réintroduire des différences entre les individus et donc de nous extirper de cette situation.

Il est quand même utile de garder à l'esprit que ceci n'est pas une solution "miracle" et qu'il est bien entendu plus intelligent de ne pas utiliser de méthodes de sélection connues pour entraîner ce type de problème.

- La mutation permet d'atteindre la propriété d'*ergodicité*.

L'ergodicité est une propriété garantissant que chaque point de l'espace de recherche puisse être atteint.

En effet, une mutation pouvant intervenir de manière aléatoire au niveau de n'importe quel locus, on a la certitude mathématique que n'importe quel permutation de notre chaîne de bits peut apparaître au sein de la population et donc que tout point de l'espace de recherche peut être atteint.

Grâce à cette propriété on est donc sûr de pouvoir atteindre l'optimum global [LA91].

II.4.6 Création De La Génération Suivante Et Critères D'arrêt

Pour réaliser une boucle d'une itération de l'AG, ce dernier doit d'abord regrouper les individus survivants après mutation dans une nouvelle population. Ensuite, l'AG va calculer la performance pour chaque nouvel individu. Enfin, si un critère d'arrêt de l'algorithme n'est pas encore atteint, la nouvelle population doit alors remplacer la population actuelle et une nouvelle boucle sera ainsi lancée.

Puisque les AGs sont des méthodes de recherches stochastiques, il est difficile de spécifier de façon rigoureuse des critères de convergence. Par exemple, la performance d'une population peut rester stable pour un certain nombre de générations avant qu'un individu supérieur puisse apparaître. Ainsi, l'application d'un critère d'arrêt devient une vraie problématique. Une pratique commune est d'arrêter l'AG après certain nombre de générations et d'examiner ensuite la qualité de la solution trouvée par rapport à la définition du problème. Un nombre typique de générations peut aller de 50 jusqu'à 500 générations.

Par ailleurs, d'autres critères peuvent être appliqués pour déterminer l'arrêt de l'AG tels que :

- l'amélioration de la solution ne dépasse plus un certain seuil
- la fonction objective du problème atteint une valeur donnée
- le temps de calcul atteint une valeur prédéterminée[HK08].

II.5 Les Apporte Des AGs Par Rapport Aux Autres Paradigmes :

- On utilise un codage des informations : on représente toutes les caractéristiques d'une solution par un ensemble de gènes, c'est-à-dire un chromosome, sous un certain codage (binaire, réel, code de Gray, etc. ...), valeurs qu'on concatène pour obtenir une chaîne de caractères qui est spécifique à une solution bien particulière (il y a une bijection entre la solution et sa représentation codée)
- On traite une population "d'individus", de solutions : cela introduit donc du parallélisme.
- L'évaluation de l'optimalité du système n'est pas dépendante vis-à-vis du domaine.
- On utilise des règles probabilistes : il n'y a pas d'énumération de l'espace de recherche, on en explore une certaine partie en étant guidé par un semi-hasard : en effet des opérateurs comme la fonction d'évaluation permet de choisir de s'intéresser à une solution qui semble

représenter un optimum local, on fait donc un choix délibéré, puis de la croiser avec une autre solution optimale localement, en général la solution obtenue par croisement est meilleure ou du même niveau que ses parents, mais ce n'est pas assuré, cela dépend des aléas du hasard, et cela est d'autant plus vrai pour l'opérateur de mutation qui ne s'applique qu'avec une certaine probabilité et dans le cas où il s'applique choisit aléatoirement sur quel(s) locus(loi) introduire des modifications[LA91].

II.6 Les Domaines D'utilisation Des Algorithmes Génétiques

II.6.1 Applications Dans La Recherche

Le problème du voyageur de commerce : Ce problème est un classique d'algorithmique. Son sujet concerne les trajets d'un voyageur de commerce. Celui-ci dispose de plusieurs points où s'arrêter et le but de l'algorithme est d'optimiser le trajet de façon à ce que celui-ci soit le plus court possible. Dans le cas où huit points d'arrêts existent, cela est encore possible par énumération (2520 possibilités [pour n arrêts, n supérieur ou égal à 3, il y a $(n-1)!/2$ chemins possibles) mais ensuite, l'augmentation du nombre d'arrêts fait suivre au nombre de possibilités une croissance exponentielle.

Par le biais d'algorithmes génétiques, il est possible de trouver des chemins relativement corrects. De plus, ce type de problèmes est assez facile à coder sous forme d'algorithme génétique. L'idée de base est de prendre comme fonction d'adaptation d'un chemin sa longueur[3]. Pour effectuer le croisement de deux chemins:

Chemin	Codage
A	1 2 3 4 : 5 6 7 8 9
D	4 1 6 3 : 9 8 2 5 7
fil	1 2 3 4 : 6 9 8 5 7

Tableau 2.1 : voyageur de commerce.

II.6.2 Applications industrielles

Un premier exemple est une réalisation effectuée au sein de l'entreprise Motorola. Le problème pour lequel Motorola a utilisé les algorithmes génétiques concerne les tests des applications informatiques. En effet, lors de chaque changement apporté à une application, il convient de retester l'application afin de voir si les modifications apportées n'ont pas eu d'influence négative sur le reste de l'application. Pour cela, la méthode classique est de définir manuellement des plans de test permettant un passage dans toutes les fonctions de l'application. Mais ce type de test nécessite un important travail humain. Le but de Motorola a donc été d'automatiser cette phase de définition de plans de tests. Ils ont pour cela défini un algorithme où chaque individu correspond à un résultat d'exécution d'un programme (l'enchaînement des valeurs passées au programme) et où chaque individu reçoit une valeur qui correspond à son aptitude à passer dans un maximum de parties du code de l'application. Finalement, l'outil développé permet, à l'aide d'un algorithme génétique, de faire évoluer ces programmes de test pour maximiser les zones testées de façon à ce que lors de modifications apportées à l'application on puisse soumettre celle-ci à des tests efficaces.

D'autres domaines industriels utilisent aujourd'hui les algorithmes génétiques. On peut retenir entre autres l'aérodynamique où des optimisations sont mises au point à l'aide de ces outils, l'optimisation structurelle, qui consiste à minimiser le poids d'une structure en tenant compte des contraintes de tension admissibles pour les différents éléments, et la recherche d'itinéraires : ces algorithmes ont été utilisés par la NASA pour la mission d'exploration de Mars, dans la gestion des déplacements du robot Pathfinder.

La société Sony les a aussi utilisés dans son robot Aibo. En effet, ce robot a « appris » à marcher dans un dispositif expérimental où son système de commande a été soumis à une évolution artificielle. Différents modes de commandes ont été testés, les plus performants ont été croisés et le résultat a été très positif. De génération en génération, le robot s'est redressé, puis a commencé à marcher en chutant souvent et a fini par marcher d'un pas assuré [3].

II.6.3 Informatique Décisionnelle

Les algorithmes génétiques sont mis en œuvre dans certains outils d'informatique décisionnelle ou de data mining par exemple pour rechercher une solution d'optimum à un problème par mutation des attributs (des variables) de la population étudiée.

Ils sont utilisés par exemple dans une étude d'optimisation d'un réseau de points de vente ou d'agences (banque, assurance, ...)[3].

II.7 Conclusion :

Dans ce chapitre on a essayé de comprendre le fonctionnement des algorithmes génétique, pour ce la on a commencer par une analogie avec la biologie pour déterminé les outils issu et utiliser dans ces dernier après on a procéder par des définitions et principes des algorithmes génétique et leur fonctionnements, ainsi que leur étapes (codage , évaluation , sélection , croisement , mutation),en détailions chaque étape tout en utilisons des exemples on a même cité les critères d'arrêt des AGs et les différents domaines d'application .

On sait que les applications des algorithmes génétiques sont multiples : optimisation de fonctions numériques difficiles, optimisation d'emplois du temps, contrôle de systèmes industriels cryptographie, apprentissage des réseaux de neurones, etc.Ils sont même utiliser dans le traitement d'images spécialement a la détection de contour qu'on va détailler dans le chapitre suivant (3eme chapitre) de notre mémoire.

III.1 Introduction

Les algorithmes génétiques ont été développés en premier par John Holland à l'université du Michigan [HO75]. Ce sont des algorithmes d'exploration fondés sur les mécanismes de la sélection naturelle et de la génétique. Ils utilisent à la fois les principes de la survie des structures les mieux adaptées, et les échanges d'information pseudo-aléatoires, pour former un algorithme d'exploration qui possède certaines des caractéristiques de l'exploration humaine. Ils sont une transcription directe du processus naturel d'évolution au monde de l'informatique.

Dans les années 1950 plusieurs chercheurs ont voulu simuler les mécanismes d'évolution génétique pour les appliquer à l'ingénierie. Mais ces travaux n'ont pas été probants en raison des connaissances insuffisantes de la génétique naturelle et aussi en raison des faibles performances des calculateurs disponibles de cette époque. Mais depuis le début des années 60, la quantité de travaux sur les algorithmes évolutionnaires suit une croissance exponentielle [SR04]

En effet, les algorithmes génétiques (AGs) ont des propriétés tels que des candidats de choix pour résoudre des problèmes d'optimisation dans lesquels la taille de l'espace de recherche est importante, où les paramètres interagissent de manière complexe et où très peu d'informations sur la fonction à optimiser sont disponibles. Ces propriétés ont motivé leur application en vision et traitement d'images.

Dans ce chapitre nous présentons les travaux effectués en vision et plus particulièrement avec les contours actifs, ensuite, nous présentons en détails l'application des contours actifs génétique (CAG) qui ont été proposé par Jean-Jacques Rousselle [ROU03] pour la détection des contours dans l'image.

II.2 Imagerie Et Algorithmes Génétiques

Ayant été reconnu comme une approche valide des problèmes nécessitant une exploration performante, et économique du point de vue calcul, les AGs sont maintenant appliqués plus largement aux domaines d'imagerie, plus particulièrement avec les contours actifs, plusieurs travaux ont été présentés dans la littérature, dans ce qui suit, nous citons quelques uns.

Une méthode de segmentation est proposée [BL99] en utilisant les contours actifs génétiques CAG (genetic snakes), qui sont un modèle actif de contours associé à une procédure de minimisation de l'énergie basée sur les AGs. Les Contours actifs génétiques (CAG) sont

proposés afin de remédier à certaines limites des contours actifs classiques, telles que : l'initialisation, l'existence des multiples minima et la sélection des paramètres d'élasticité. Ballerini a prolongé la formulation des CAG de deux manières, en explorant le terme interne additionnel et le terme correspondant à l'énergie externe, des applications sur des images couleurs ont été données. Aussi, [ROU03] qui a proposé une implémentation originale basée sur les algorithmes génétiques pour détecter dans une image d'articulation fémoro-tibiale la partie bombée du fémur « le condyle ».

III.3 Les Contours Actifs Génétique

III.3.1 Origine

Il s'agit, pour nous, de mettre la capacité exploratoire des algorithmes génétiques au service de la recherche d'un contour que nous définissons à partir de la méthode des contours actifs. Ainsi, la solution recherchée est un contour, c'est-à-dire un ensemble de points et la fonction d'évaluation est liée à l'énergie totale du *snake*

La différence avec la plupart des algorithmes de contours actifs réside dans le fait que l'on ne fait pas évoluer un *snake* mais une population de *snakes*. Seul le meilleur *snake* de la population à un instant donné est considéré comme un l'approximation du contour. Si la taille de la population est suffisamment élevée, on peut se permettre d'initialiser les contours de manière totalement aléatoire dans l'image.

C'est à l'aide des opérateurs de croisement et de mutation qu'on assure la variation de la population ainsi la convergence vers des individus plus intéressants. Cela permet de s'affranchir de toute connaissance structurelle a priori de l'image.

Les snakes ainsi traités ne se « déplacent » pas à proprement parler sous l'effet des forces mises en jeu, mais sont recombinaisonnés (croisement) entre eux et mutés afin d'en créer d'autres, d'énergie moindre, qui ont atteint une bonne stabilité en fonction des forces mises en jeu. Ainsi, la notion de déplacement dans le « voisinage » de chaque point, disparaît totalement [ROU03].

III.3.2 Implémentation Des Contours Actifs Génétiques

Pour réussir une telle configuration il faut comprendre les principes de bases à la fois de contours actifs (snakes) et les algorithmes génétiques pour savoir les combiner entre ces deux approches, les étapes de l'implémentation sont presque les mêmes que les étapes d'un algorithme génétique. Comme on a détaillé dans le 2ème chapitre (II.4) :

Chapitre 03 : Les Contours Actifs Génétiques

III.3.2.1 Codage Des Chromosomes

Le codage est la manière dont la réalité physique du snake(phénotype) est traduite en une chaîne binaire (génotype).

Nous allons envisager deux approches qui reposent respectivement sur les coordonnées cartésiennes et sur les coordonnées polaires.

A- Coordonnées Cartésiennes

Le contour étant un ensemble de points, le codage intuitif consiste à concaténer les coordonnées cartésiennes de chaque point.

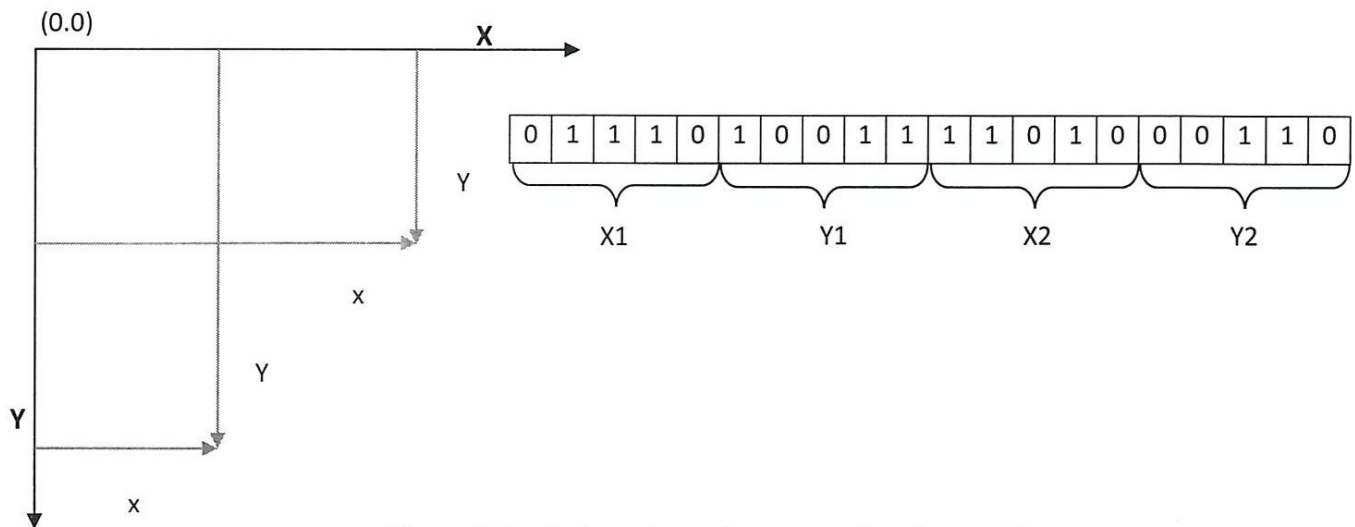


Figure 3.1 : Codage du snake en coordonnées cartésiennes.

Pour un snake de 40 points et une image de 1024 x 1024 pixels, cela donne un total de 800 bits, car 10 bits sont nécessaires pour écrire une coordonnée en binaire.

Un codage des coordonnées sur 10 bits, introduit un problème de « contrainte » fréquent dans le domaine des algorithmes génétiques. Si l'image est de taille inférieure à 1024 pixels, certaines chaînes pourront représenter des solutions non valides, des suites de points hors image.

Pour remédier à ce problème on effectue l'un des traitements suivants :

Au niveau de la représentation : en interdisant à tous les points de l'espace de recherche de coder une solution non réalisable.

Au niveau des opérateurs génétiques : dans ce cas, il peut exister des points de l'espace de recherche qui ne soient plus des solutions réalisables. Les opérateurs génétiques sont alors conçus de manière à ce qu'ils ne génèrent jamais l'un de ces points.

Avec un opérateur de réparation : celui-ci va modifier légèrement les individus non réalisables pour les rendre réalisables.

Au niveau de la fonction d'évaluation : si aucune des techniques précédentes ne peut être employée, il est possible d'infliger aux individus non réalisables une pénalité importante en matière de fitness afin qu'ils soient éliminés par la pression sélective [MZ92].

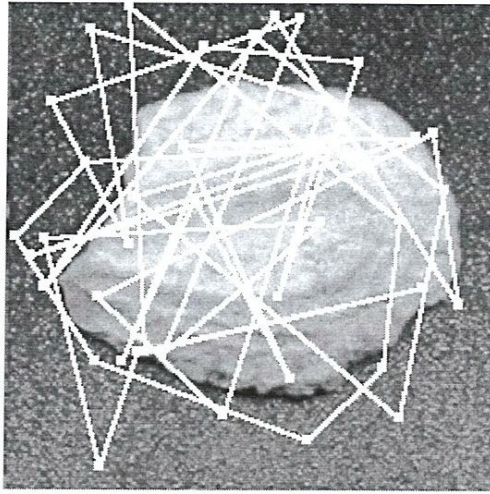


Figure 3.2 : Initialisation des points en coordonnées cartésiennes sans réordonner les points .

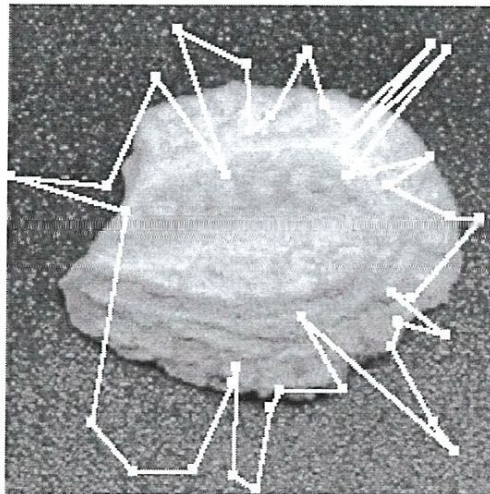


Figure 3.3 : Initialisation des points en coordonnées cartésiennes en réordonnant les points.

L'aspect ordonné des points du snake est traduit pour chaque point par une place au sein du chromosome. Cette convention conduit à différents problèmes, d'une part au niveau de l'initialisation de la population de snakes, et d'autre part au niveau de l'application des opérateurs.

La première étape de l'algorithme génétique est la construction d'une population initiale, un ensemble de chromosomes qui représentent des contours. Ces chromosomes doivent être déterminés aléatoirement, c'est-à-dire que pour chaque contour, est choisie aléatoirement une

suite de points. Ces points sont ordonnés par l'ordre du tirage. Les suites de points présentent des boucles comme sur la Figure(3.3). Le contour n'est pas dans les conditions d'une initialisation classique où les points forment déjà un contour fermé. Cela signifie qu'en plus de minimiser l'énergie globale l'algorithme doit résoudre un problème analogue à celui du voyageur de commerce qui n'est pas souhaitable. Il convient alors de réordonner les points, soit en modifiant directement les chromosomes de la population initiale (ce qui revient à faire une initialisation semi-aléatoire), soit à partir de la fonction d'évaluation, juste après le décodage du génome.

Dans une image, l'objet d'intérêt est généralement relativement centré. Pour réordonner les points, nous avons calculé leurs coordonnées polaires (p, θ) ayant pour origine le centre (x_c, y_c) de l'image.

Il suffit ensuite de trier les points suivant les θ croissants comme on peut le voir sur la Figure (3,4).

Les formules de conversion sont les suivantes :

$$p = \sqrt{(x-x_c)^2 + (y-y_c)^2} \quad (\text{III.1})$$

$$\theta = \begin{cases} \arccos(\mu) & \text{si } \mu \geq 0 \\ \arcsin(\vartheta) & \text{si } \mu < 0 \text{ et } \vartheta \geq 0 \\ -\arccos(\mu) & \text{si } \mu < 0 \text{ et } \vartheta < 0 \end{cases} \quad (\text{III.2})$$

$$\text{Avec } \mu = \frac{x-x_c}{p} \text{ et } \vartheta = \frac{y-y_c}{p} \text{ (dans ce cas } \theta \in [-\pi, \pi]) \quad (\text{III.3})$$

$$x = x_c + p \cos \theta \text{ et } y = y_c - p \sin \theta \quad (\text{III.4})$$

Il est évident que la qualité de la réorganisation dépend de la forme du contour recherché. Sachant que cette méthode ne serait alors applicable qu'aux objets convexes étoilés, nous pouvons retenir une solution ayant recours aux coordonnées polaires[4].

B- Coordonnées Polaires

Comme nous utilisons les coordonnées polaires pour trier les points, nous avons choisi de ne pas passer par les coordonnées cartésiennes, mais nous avons opté pour un codage directement en polaire des points dans le chromosome. Dans ce cas, l'angle θ variera de 0 à 2π et pour une image de largeur L et de hauteur H l'espace de définition devient : P_{MAX}

$$P - MAX = \frac{L}{2 \cos \theta} \text{ si } \theta \in [0, \frac{\pi}{4}] \cup [\frac{7}{4}\pi, 2\pi] \quad (\text{III.5})$$

$$P - MAX = \frac{H}{2 \sin \theta} \text{ si } \theta \in [\frac{\pi}{4}, 3\frac{\pi}{4}] \quad (\text{III.6})$$

$$P - MAX = \frac{-L}{2 \cos \theta} \text{ si } \theta \in [3\frac{\pi}{4}, 5\frac{\pi}{4}] \quad (\text{III.7})$$

$$P - MAX = \frac{-H}{2 \cos \theta} \text{ si } \theta \in [5\frac{\pi}{4}, 7\frac{\pi}{4}] \quad (\text{III.8})$$

Afin de réduire l'espace de recherche, sachant que nous cherchons un contour fermé plus ou moins centré sur l'image, nous rendons implicite la valeur de θ . Les points sont répartis dans toutes les directions à partir de l'origine, c'est-à-dire pour un snake de 40 points, on prend $V_i = (p_i, \theta)$ avec $\theta = \frac{i}{20}\pi$. Dans ce codage, seul p est explicitement codé ; cela réduit de moitié la longueur des chromosomes. L'espace de recherche n'est donc plus le plan de l'image mais l'union des 40 axes concourants [ROU03].

III.3.2.2 Sélection

Comme nous l'avons déjà précisé, plusieurs méthodes de sélection sont possibles, dont la roue biaisée, la sélection universelle stochastique et élitiste et le K-tournoi précédemment vue dans le 2ème chapitre ces méthodes sont peu robustes et nécessitent de trier tous les chromosomes de la population selon leur fitness. Cette opération est coûteuse en temps de calcul. Nous avons limité notre étude aux méthodes de sélection en K-tournois, en faisant varier le paramètre K.

Pour une population de taille N, la méthode peut être décrite par trois étapes.

1. Choisir K individus parmi les N présents dans la population.
2. Recopier le meilleur de ces K individus dans la nouvelle population.
3. Recommencer l'opération N fois.

Un paramètre K élevé a tendance à accroître la pression sélective, tandis qu'une valeur de K faible la diminue. Une forte pression sélective se traduit par une faible « espérance de vie » des chromosomes médiocres. En d'autres termes, lors de l'exploration de l'espace de recherche, une faible valeur de K privilégie la diversité génétique, tandis qu'une forte valeur de K privilégie l'exploitation des meilleurs schémas trouvés, au risque de tomber dans un optimum local.

La Figure (3.4) présente un test de cette méthode avec différentes valeurs de K. Chaque courbe représente la moyenne sur 25 exécutions de l'algorithme pour une valeur de K.

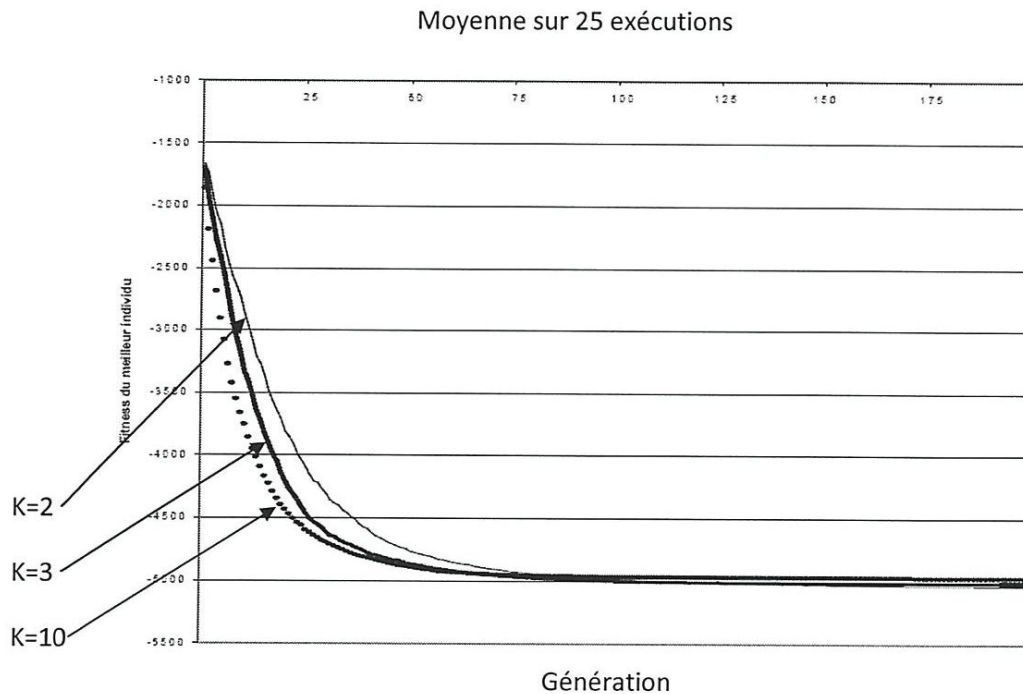


Figure 3.4 : les tests de la Sélection par k-tournoi.

Nous pouvons remarquer que les différences sont faibles mais qu'avec un K élevé, la convergence est plutôt plus rapide mais la solution finale trouvée est de qualité moyenne, tandis qu'avec un K faible, la convergence est plus lente, mais permet d'aboutir à une solution plutôt de meilleure qualité. Le meilleur compromis exploration/exploitation semble être atteint avec $K=3$ [ROU03].

III.3.2.3 Croisement

L'utilisation d'une représentation binaire du génome limite notre choix de méthodes de croisements. Nous avons testé les croisements à un ou plusieurs points de coupure, ainsi que le croisement uniforme (voir II.4.2)

L'expérimentation montre que le choix d'une méthode de croisement affecte peu la qualité de la solution finale obtenue. En revanche, cela a une forte influence sur la vitesse de convergence de l'algorithme. Nous avons donc étudié l'évolution de la fitness sur les 50 premières générations.

On peut observer sur la Figure (3.5) que plus le nombre de points de coupure est élevé, plus la convergence semble rapide. Quant aux résultats obtenus par croisement uniforme, ils se situent bien en deçà de ceux fournis par un croisement à un point de coupure.

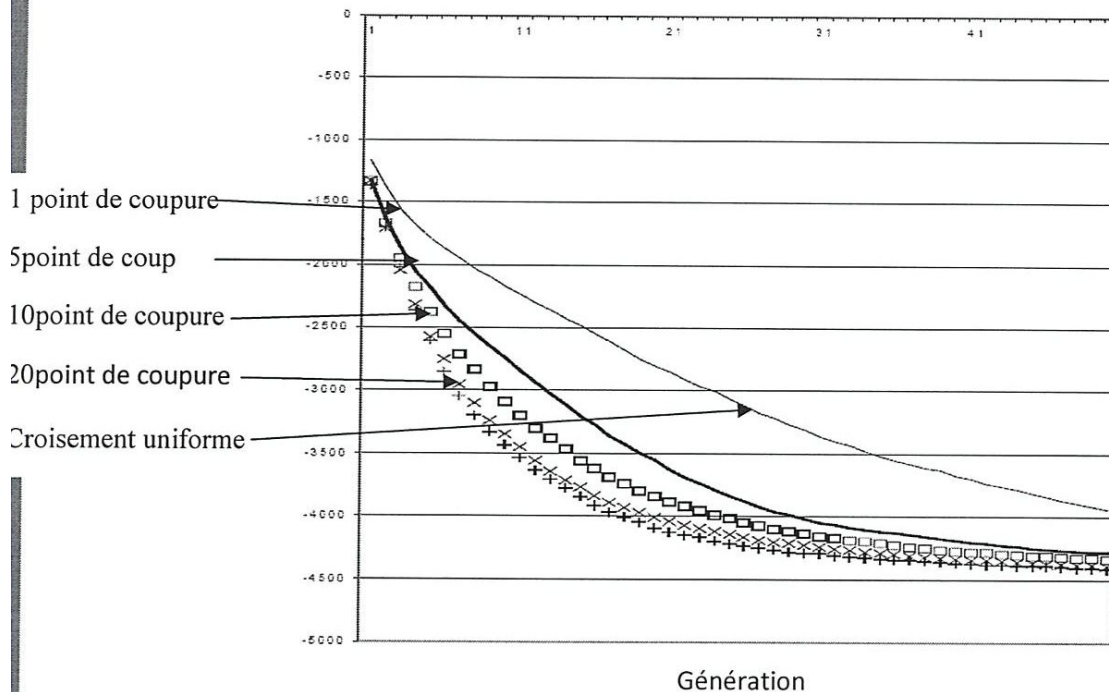


Figure 3.5 : comparaison entre les différents types de croisement.

Nous pouvons constater également qu'au-delà de dix points de coupure, la vitesse de convergence augmente peu. Nous avons donc retenu cette méthode qui présente l'avantage de fournir des résultats très corrects sans pour autant être trop coûteuse en temps de calcul [ROU03].

III.3.2.4 Mutation

Le choix du taux de mutation a une grande influence sur l'efficacité de l'algorithme. Traditionnellement, pour les premiers essais, le taux de mutation μ est fixé à 0,001. Dans un second temps, la formule $\mu = \frac{1}{L}$ (où L est la longueur d'un chromosome) est fréquemment utilisée. Ce taux garantit qu'en moyenne, un gène par chromosome sera affecté. Cependant, il n'y a pas de règle précise pour déterminer le taux de mutation idéal, mais on peut mentionner quelques remarques de logique qu'il convient de respecter.

La mutation permet (entre autres) de palier le manque de diversité de la population initiale. En d'autres termes plus la population initiale est grande plus la mutation peut être petite

Concèderons que les chromosomes ont une longueur de 440 bits, donc le taux de mutation « Idéal » devrait se situer aux alentours de 0,002. Les Figure (3.8)et Figure (3.9) présentent les résultats obtenus sur une même image en faisant varier μ de 0,0005 à 0,01. Chaque courbe représente la moyenne des résultats obtenus sur 25 exécutions de l'algorithme.

Par ailleurs, le taux de mutation a également une influence sur la régularité de la convergence, et le fait d'effectuer une moyenne sur plusieurs exécutions de l'algorithme atténue la visibilité de ce phénomène. C'est pourquoi nous présentons aussi les résultats obtenus durant les dernières itérations, sur une seule exécution [BT93].

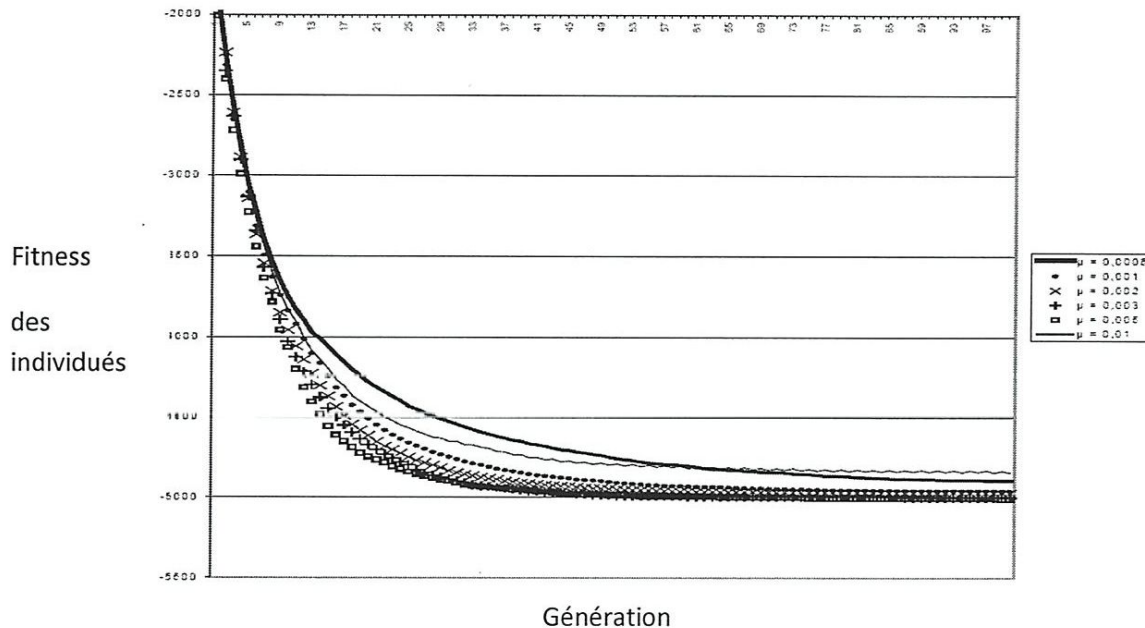


Figure 3.6: les taux de mutation (moyenne sur 25 exécutions).

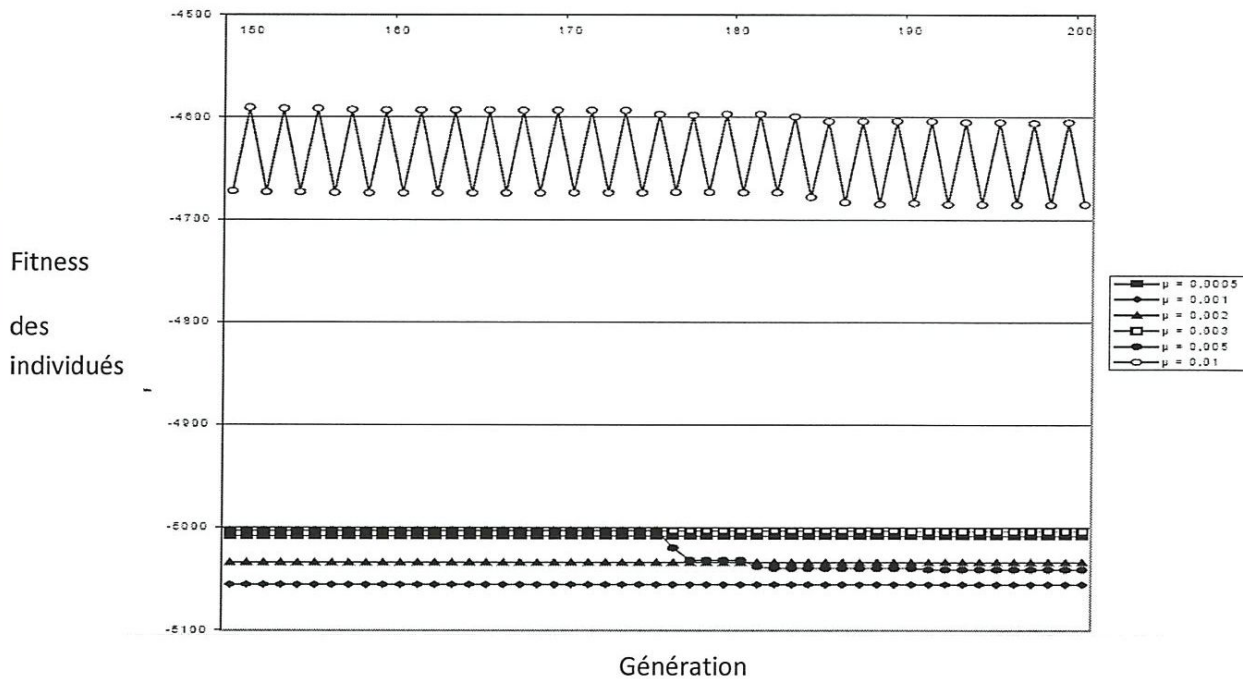


Figure 3.7: Influence de la mutation, zoom sur les générations 150 à 200 (une exécution).

Comme nous pouvons l'observer sur le premier graphique (Figure 3.7), jusqu'à un certain seuil (0,005), plus le taux de mutation est élevé, plus l'algorithme converge rapidement. Lorsque ce seuil est dépassé, la convergence est ralentie par la mutation qui, lorsqu'elle est trop élevée, détériore systématiquement les chromosomes de bonne qualité issus du croisement.

Lorsque μ vaut moins de 0,001, la convergence est très régulière mais lente, et la solution finale est de qualité moyenne par rapport aux solutions obtenues avec des valeurs supérieures.

D'après Jean Jacques ROUSSELLE[ROU03] il avait constaté qu'il était courant d'obtenir un snake globalement très bon mais présentant un ou deux points anguleux que nous retrouvons dans tous les chromosomes de la population (effet dû à la convergence), probablement parce qu'ils sont apparus pour la première fois dans un snake de bonne qualité qui est devenu omniprésent au fil des générations. Lorsque l'algorithme a longtemps convergé, tous les chromosomes sont pratiquement identiques (uniforme), et seule la mutation peut faire apparaître de nouveaux schémas.

Ce phénomène d'uniformisation de la population est illustré sur la Figure (3.8). Chaque ligne représente un chromosome et chaque colonne un allèle. Dans ce type de représentation (utilisé dans [Dessalles 1996]), l'apparition de structures en forme de « codes barres » traduit une perte de diversité au sein de la population.

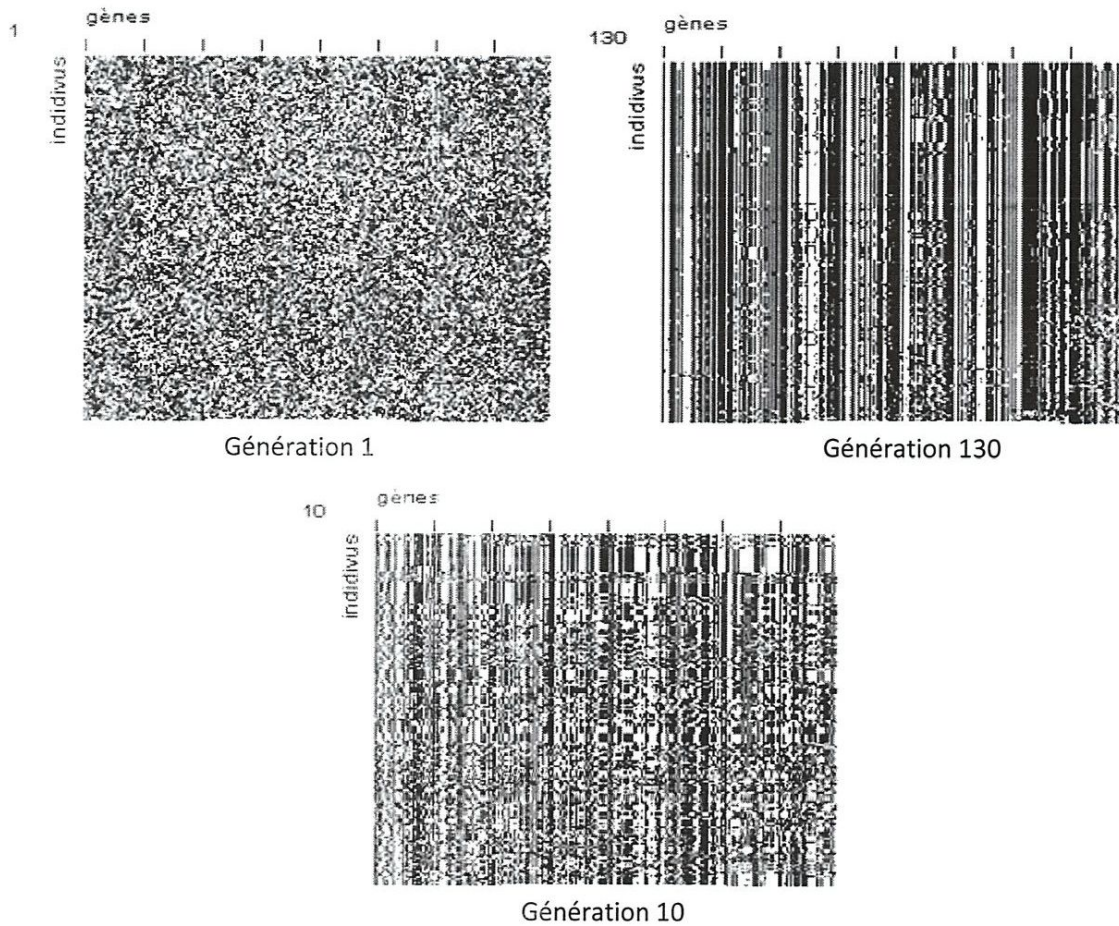


Figure 3.8 : uniformisation de la population.

Dans ces conditions, il pourrait être intéressant d'augmenter le taux de mutation uniquement sur les points considérés comme « mauvais » par rapport aux autres dans ce cas, la mutation est dite adaptative.

Nous avons réalisé ce procédé en insérant un bit de contrôle entre chaque groupe de bits codant un point. Le choix de la valeur de ce bit sera effectué par la fonction d'évaluation. Nous avons mis ce bit à 1 si le point est mauvais, et à 0 dans le cas contraire. Ce sera ensuite l'opérateur de mutation qui, en fonction de la valeur de ce bit, appliquera soit le taux de mutation normal, soit un taux plus élevé, sur les gènes concernés [DJ96].

III.4 Conclusion :

Dans ce chapitre, nous venons de rappeler les principes d'un outil de résolution qui a fait ses preuves en I.A, à savoir les algorithmes génétiques. Leur principe, basé sur une description fidèle d'une évaluation naturelle, assure une recherche efficace dans le monde des solutions d'un problème donné. Lorsque leurs paramètres (taille de la population, nombre d'itérations, probabilités de tirages,...) sont bien choisis, ainsi que le codage des données est encore plus important dans le processus d'exploration. Il n'y a aucun jeu de paramètres qui est universel pour tous les problèmes, c'est à dire que ces valeurs dépendent étroitement du type de problème à résoudre. Un autre inconvénient est que pour le moment, on ne sait pas prédire qu'un algorithme génétique convergera vers la bonne solution.

Les algorithmes génétiques ont prouvé son efficacité pour la résolution de certains problèmes d'optimisation complexes. Et beaucoup de domaines d'applications ont utilisés les AGs. Nous avons présenté quelques travaux effectués dans des applications de la vision et de la détection des contours.

Les algorithmes génétiques ont été appliqués à la détection des contours pour la recherche d'un contour optimale dans l'espace des contours possibles, et pour optimiser les paramètres d'un algorithme de détection des contours existant.

Dans tous les cas, il est nécessaire de définir une fonction de fitness qui calcule la qualité de contour actif. Cette fonction de fitness dépend énormément au problème à résoudre, et il n'existe pas une mesure régulièrement admise. Il est également nécessaire de déterminer les valeurs des paramètres (codification des individus, mode de sélection, mode de croisement, taux de mutation ...) qui gèrent le comportement de ces algorithmes, il est souvent nécessaire d'ajusté les paramètres des algorithmes génétiques empiriquement.

Nous avons utilisé cette méthode pour la détection des contours des zones suspectes dans les images mammographique ainsi que l'extraction des paramètres liés à la masse surtout en ce qui concerne la forme, la taille et les contours des lésions.

l'avantage de cette méthode est que tous les points de l'espace de recherche peuvent être évalués à chaque itération, cette technique permet de détecter un contour qui n'est pas bien marqué. Le problème que nous nous sommes proposé de résoudre est un problème difficile à cause de la nature des données à traité d'une part (images trop bruitées) et de la nature de l'application médicale d'autre part.

Chapitre 04 : Réalisation & Conception

VI.1 Introduction

L'analyse d'image est une science appliquée, développée pour résoudre des problèmes de vision. A l'intérieur des domaines privilégiés de l'analyse d'image, la détection des contours a pris une part tout à fait originale, grâce à son approche aussi bien pragmatique que théoriquement bien fondée. C'est sans doute grâce à la symbiose entre une rigueur mathématique et une volonté d'appliquer ses principes à des vrais problèmes que les méthodes de détection des contours ont connu le succès qu'elles méritent, auprès, en particulier, des industriels et imagerie médicale.

Nous présentons dans ce chapitre l'application relative à l'analyse des mammographies. L'analyse automatique des images mammographiques est un problème qui n'a été abordé qu'assez récemment et vient en parallèle avec un travail de développement de techniques de mammographies numériques. Ce domaine de recherche est aujourd'hui en plein essor.

Cette application fait appel aux notions que nous avons présentées dans ce mémoire ainsi qu'à d'autres notions plus classiques de détection des contours, filtrage et segmentation. Nous indiquons dans ce chapitre la démarche générale adoptée pour le processus de l'extraction automatique des paramètres liés aux masses dans les clichés mammaire. Ensuite, nous montrons et discutons les résultats obtenus lors de la procédure d'extraction.

VI.2 Environnement De Développement

VI.2.1 Besoins Matériels

Pour le développement de notre logiciel, on a travaillé sur un micro-ordinateur ayant les caractéristiques suivantes :

- ✦ Processeur : Intel®Core™2 Duo CPU 2.93 GHZ
- ✦ RAM : 4 GB
- ✦ Disque dur : 500 Go

VI.2.2 Besoins Logiciels

L'environnement de développement qu'on a choisi pour l'implémentation de notre projet est « Eclipse » qui est principalement écrit en Java (à l'aide de la bibliothèque graphique SWT (StandardWidgetToolkit), d'IBM (International Business Machines)). En effet, ce choix n'a pas été aléatoire, JAVA est un langage simple, orienté objets, interprété, fiable, sécurisé, indépendant de toute architecture, portable, de haute performance, multithread.

- **La Simplicité De Java**

Java est un système qui puisse être programmé simplement sans nécessiter un apprentissage ésotérique, et qui tire parti de l'expérience standard actuelle, Java a été conçu de faciliter la compréhension du système. De plus, pour épargner à l'utilisateur la difficulté de gestion de mémoire, Java comprend un gestionnaire de mémoire automatique.

- **Orienté Objet**

Pour rester simples, disons que la conception orientée objet est une technique de programmation qui se concentre sur les données (les objets) et sur les interfaces avec ces objets. Pour faire une analogie avec la menuiserie, on pourrait dire qu'un menuisier "orienté objet" s'intéresse essentiellement à la chaise l'objet qu'il fabrique et non à sa conception (le "comment").

- **Interprété**

L'interpréteur Java peut exécuter les bytecode directement sur n'importe quelle machine sur laquelle il a été porté. Dans la mesure où la liaison est un processus plus incrémentiel et léger, le processus de développement peut se révéler plus rapide et exploratoire.

- **Fiabilité**

Java a été conçue pour que les programmes qui l'utilisent soient fiables sous différents aspects. Sa conception encourage le programmeur à traquer préventivement les éventuels problèmes, à lancer des vérifications dynamiques en cours d'exécution et à éliminer les situations génératrices d'erreurs.

- **Sécurité**

Java a été conçue pour être exploitée dans des environnements serveur et distribués. Dans ce but, la sécurité n'a pas été négligée. Java permet la construction de systèmes inaltérables et sans virus.

- **Portable**

On ne trouve pas les aspects de dépendance de la mise en oeuvre dans la spécification. Les tailles des types de données primaires sont spécifiées, ainsi que le comportement arithmétique qui leur est applicable.

- **Performances Elevées**

En général, les performances des bytecodes interprétés sont tout à fait suffisantes, il existe toutefois des situations dans lesquelles des performances plus élevées sont nécessaires. Les bytecodes peuvent être traduits à la volée en code machine pour l'unité centrale destinée à accueillir l'application.

- **Multithread**

Les avantages du multithread sont une meilleure interactivité et un meilleur comportement en temps réel.

On parlant d'Eclipse maintenant, c'est un environnement de développement intégré libre (le terme Eclipse désigne également le projet correspondant, lancé par IBM) extensible, universel et polyvalent, permettant potentiellement de créer des projets de développement mettant en œuvre n'importe quel langage de programmation, et ce langage, grâce à des bibliothèques spécifiques, est également utilisé pour écrire des extensions.

VI.2.3 La Version D'Eclipse

La version qu'on a utilisé pour notre application appelée hiliose (la version 3.6.1), parue en juin 2010.

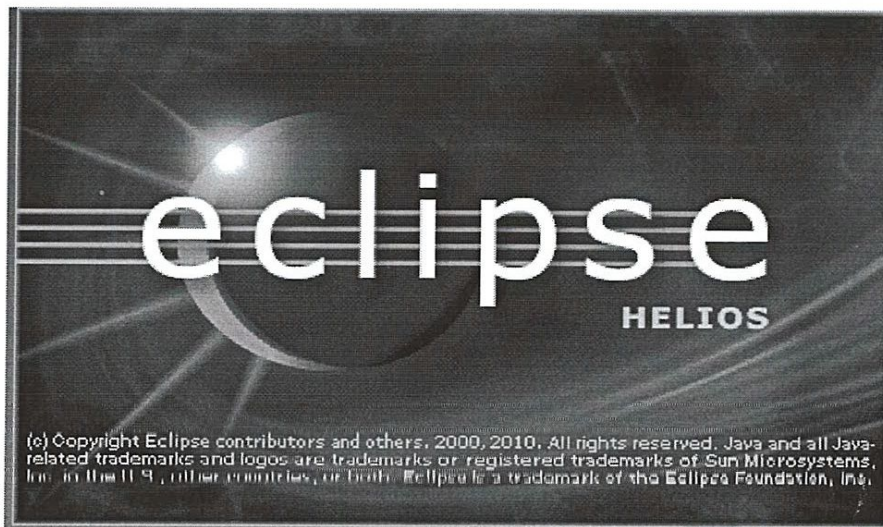


Figure 4.1 : démarrage d'eclipse.

Dans notre application on a utilisé un format d'image intitulé DICOM qui nécessite pour sa manipulation par Eclipse une importation d'une bibliothèque graphique appelé « **image j** » version 1.45s mise en place par l'institut américaine «Wayne Rashband,Nationale Institues of Health,USA » disponible gratuitement et en open source dans le site «<http://imagej.nih.gov/ij>»,qui tourne sous machine virtuelle java version 1.6.0_31(32bit), Le format de fichier image utiliscr dans notre application c'est le format DICOM.

VI.2.4 Le Format DICOM

La norme DICOM "Digital Imaging and Communication in Médecine" est un document qui définit une méthode de communication pour les différents équipements d'imagerie médicale

numérique. Cette norme est maintenant utilisée par la plupart des fabricants de matériel d'imagerie médicale. Elle représente un progrès réel car elle permet de prédire la possibilité d'interconnexion des appareils à partir d'un Document de Conformité ou " ConformanceStatement " émis pour chacune des machines respectant cette norme.

Ainsi la norme permet aux équipements de communiquer localement, à distance ou au travers d'un média en assurant la compatibilité des équipements et en éliminant les formats propriétaires.

La norme DICOM a été émise par l' ACR (American College of Radiology) en association avec la NEMA (National ElectricalManufacturers Association) Actuellement mise à jour par ces deux comités auxquelles se sont joints d'autres comités d'experts internationaux tels le JRIA au Japon, l'ANSI aux USA , le CENTC251 en Europe.

L'ACR est responsable des consignes techniques et médicales, la NEMA est responsable des publications et des problèmes légaux, pour éviter des conflits d'intérêts ou une éventuelle infraction aux lois anti-trust.

Ces comités n'ont pas d'intérêts économiques, on notera que cette norme est définie hors de France, en langue anglaise avec les difficultés de traduction d'un langage technique précis [5].

Un fichier DICOM = image + données patient en un seul fichier

Image : codées en binaire, en général non compressées

Données : nom du patient, type d'examen, hôpital, date d'examen, type d'acquisition etc...

Extension : .dc3, .dcm, .dic ou sans extension

VI.3 Conception

Lors d'un examen mammographique, le radiologue ne fait que résoudre un problème de vision. C'est à partir de ce point et de récentes techniques de l'analyse et prétraitement d'image, que l'idée d'utiliser la machine pour faciliter le travail du radiologue pour mieux diagnostiquer un cliché mammographique. La recherche menée dans ce domaine actuellement est encore à un stade peu avancé. Pour l'instant, seul le problème de la détection des microcalcifications a prouvé sa faisabilité. Les résultats plus qu'encourageants obtenus dans ce domaine ont permis d'espérer un aboutissement aux autres problèmes plus complexes comme celui des masses.

La conception de Notre application se résume dans le l'organigramme suivant :

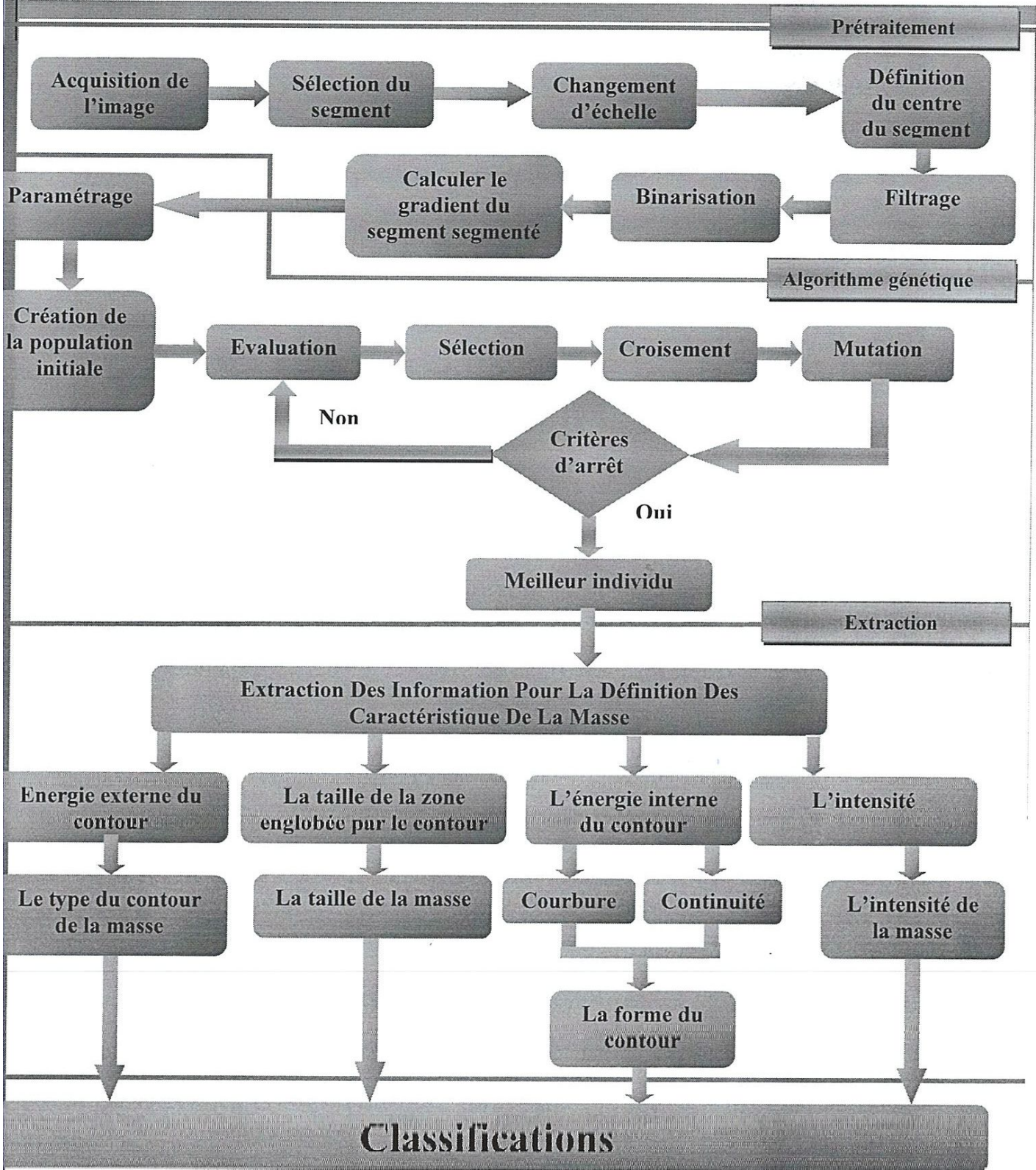


Figure 4.2 : le diagramme de conception.

Le but de notre application est d'extraire les paramètres de la masse pour procéder à la classification des masses segmentés selon le système d'aide à la rédaction des comptes rendus BI-RADS (Breast Imaging Reporting System and Data System) [LL05]. Notre conception se divise en trois parties (Figure (4.2)).

VI.3.1 Prétraitement de L'image

Après avoir chargé l'image et avant d'entamer notre méthode pour l'extraction des paramètres, des étapes de prétraitement est nécessaires car nous n'avons pas des images DICOM bien segmenté; La préparation de l'image se fait selon les étapes suivantes :

- A- **Sélection du Segment** : on détermine la zone suspecte à partir de l'image chargé sous forme d'un rectangle.
- B- **Calcul du gradient** : le gradient du segment utiliser pour calculer les énergies externes du contour actif génétique.
- C- **Changement d'échelle** : pour que les valeurs des paramètres détectés des masses soient uniforme quel que soit la taille de la masse, alors en doit changer la taille de la zone sélectionnèrent respectant une échelle de référence.
- D- **Définition du centre du segment** : C'est important de fixé le centre du segment et récupéré ses coordonné pour procédé a l'étape de codage de la population qu'on va entamer dans la suite.
- E **Filtrage** : Dans une image on peut confondre entre les traie de texture et les contours alors pour remédier a ce problème on' a appliqué un lissage (filtrage moyeneur) Sur le segment.

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \times \frac{1}{9}$$

Figure (4.3) : le filtre moyeneur.

- F- **Binarisation** : Il existe plusieurs méthodes de segmentation, le choix de la méthode dépend du contexte d'utilisation dans note cas la méthode utiliser et : la segmentation par seuillage.

Le but de notre segmentation c'est l'accentuation des contours pour maximisé le gradient et la séparation du segment en objet et fond, le fond sera coloré en noire et l'objet en blanc.

$$\left\{ \begin{array}{l} \text{Si } I(i,j) \geq S \\ \quad \text{Alors } I(i,j) = 255 \text{ (blanch)} \\ \text{Si Non} \\ \quad I(i,j) = 0 \text{ (noir)} \end{array} \right.$$

Ou « I » la matrice image, « S » le seuil, $I(i,j)$ c'est l'élément de l'image.

G- Calcul du Gradient après binarisation : Pour détecté le contour finale du segment, on 'a applique un filtre de Sobel déjà vue dans le chapitre 1 de notre mémoire (voir (I.8.3.5.a)).

-1	0	1	-1	-2	-1
-2	0	2	0	0	0
-1	0	1	1	2	1

Figure4.4 : le filtre sobel.

VI.3.2 Algorithme Génétique :

Nous utilisons la méthode des contours actifs génétiques (CAG) pour détecter les contours de l'anomalie, et l'extraction des paramètres liés à la masse.

VI.3.2.1 Paramétrage

- A- Réglage des paramètres de l'algorithme génétique :** Dans cette étape nous définissons la taille de chromosome (nombre de point dans le contour actif) et la taille de la population initial (le nombre des contours à générer).
- B- Réglage des Coefficients des Energies:** La valeur de fitness du contour actif génétique dépend énormément des coefficients des énergies de contour α, β, γ et θ

$$F = \alpha * \text{courbure} + \beta * \text{contunuité} + \gamma * \text{gradient} + \theta * \text{context} \quad (\text{VI.1})$$

Ou F est la valeur de fitness

VI.3.2.2 Initialisation De La population Initiale (Codage)

Pour codé un contour on a utilise les coordonné polaire (R, ϑ) (III.3.2.1.B).

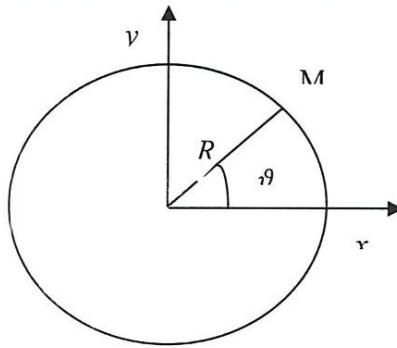


Figure (4.5) : les coordonnées polaire.

R : rayon .

ϑ : l'angle entre le rayon R et x.

M :le point a codé .

Les coordonnées de M se donne par le terme suivant : $M(R, \vartheta)$.

Pour N points on a utilisé un tableau (T) de longueur N , la figure (4,6) montre ce tableau :

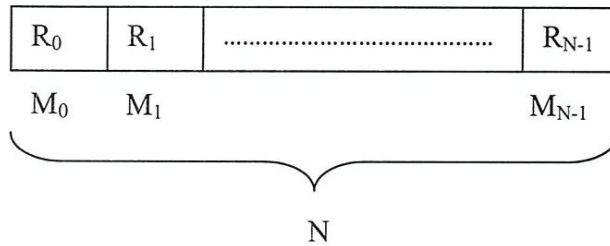


Figure (4.6) : le tableau de codage des points d'un contour.

Les coordonné sont les suivants :

$$\vartheta = \frac{Ri}{N} * \pi$$

$$M_0(R_0, \frac{0}{N} * \pi), M_1(R_1, \frac{1}{N} * \pi) \dots \dots \dots M_{N-1}(R_{N-1}, \frac{N-1}{N} * \pi) \quad (VI. 2)$$

i : l'indice du tableau (T)

Voila un pseudo code récapitulant notre codage :

```

Pour tous les contours de la population
|
| Pour chaque point d'un contour Faire :
| | Généré un rayon aléatoire R selon  $\vartheta$ (R ne dépasse pas les limite du segment).
| | Fin pour ;
| Fin pour ;

```

VI.3.2.3 Evaluation

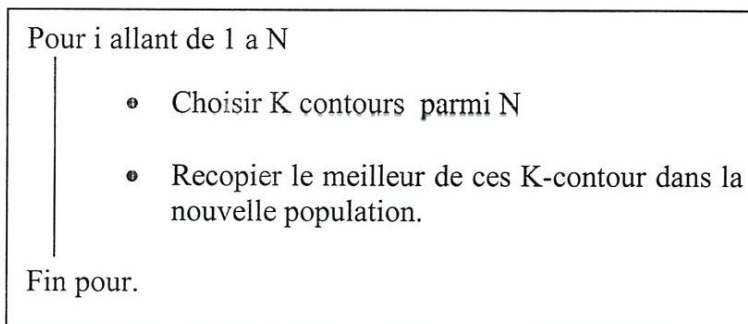
La fitness des contours actifs génétiques (CAG) est calculé par la fonction suivante :

$$F_{CAG} = \sum_{i=1}^{40} f(i) \quad (IV.3)$$

Où $f(i)$ est la fitness du chromosome i .

VI.3.2.4 Sélection

La méthode de sélection choisie est le K-tournoi présenté à la section (II.4.3.C), dans ce cas nous choisissons 3 contours parents aléatoirement parmi les N présents de la population, ensuite on recopie le meilleur parmi les trois contours dans la nouvelle population. Et on recommencer l'opération $(N/2)$ fois. C'est-à-dire que la proportion de parents qui sont remplacés par leurs descendants égale à 0.5 pour éviter le fossé des générations, de cette façon la nouvelle population aura $(N/2)$ contours qui sont recopiés directement à partir de l'ancienne génération, et $(N/2)$ contours qui vont être générés par les opérateurs de croisement et mutation sur les 100 contours sélectionnés.



Ou {

- N : c'est la taille de la population initiale.
- K : nombre aléatoire qui varier pour chaque interaction.

VI.3.2.5 Croisement

La méthode de croisement choisie consiste à combiner deux individus choisis aléatoirement parmi les 100 contours sélectionnés, pour en ressortir deux autres contours enfants, pour cela

nous avons choisi dix points de coupure aléatoirement tout au long de la courbe, de cette façon 100 contours enfants sont recopies a la nouvelle génération.

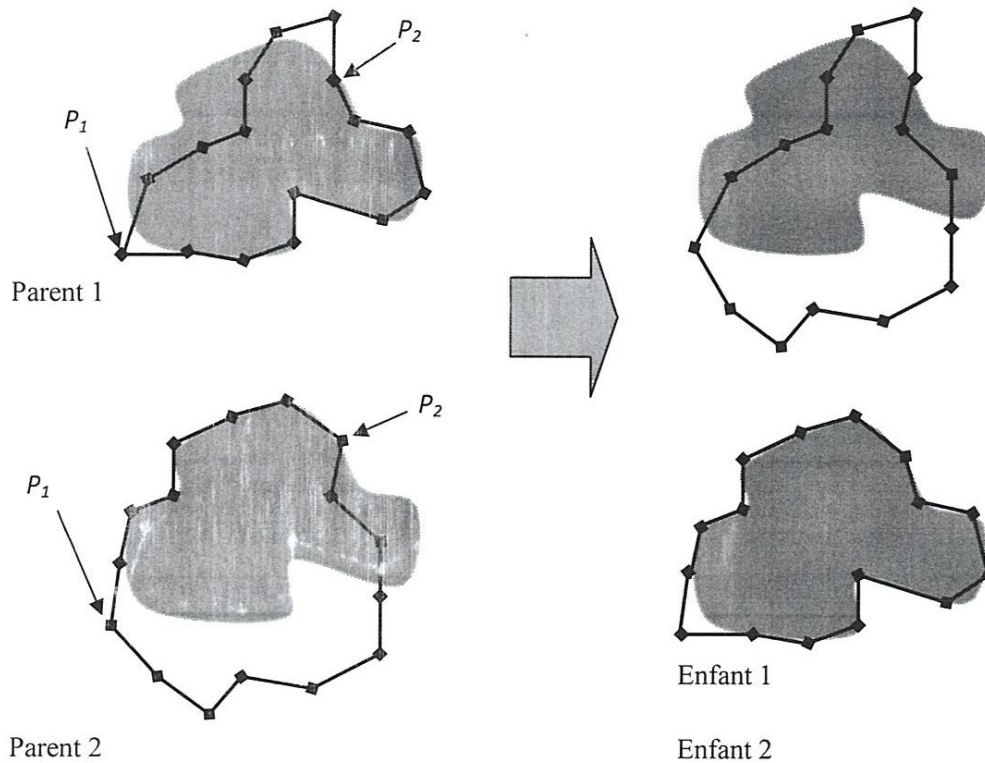


Figure (4.7) :croisement de deux contours en 2 points de coupure P_1 et P_2

VI.3.2.6 Mutation

Nous avons appliqué l'opérateur de mutation seulement sur les $(N/2)$ individus enfants, pour ce faire, nous avons optée pour une opération adaptative, c'est-à-dire que seul le mauvais chromosome dans le contour est muté, et le taux de mutation égale à 0,025.

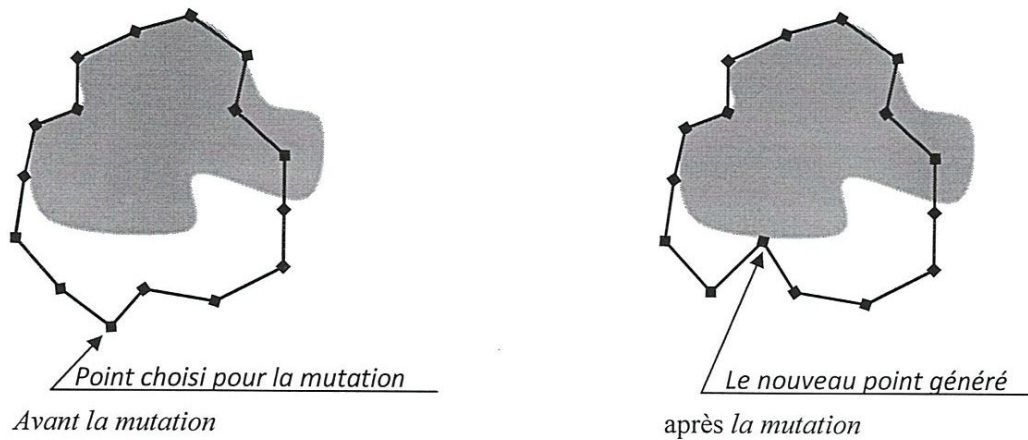


Figure (4.8) : représentation de la phase de mutation dans CAG

VI.3.2.7 Critère d'arrêt

Une fois le nombre maximum de générations atteint, fixé par l'utilisateur, nous choisissons le meilleur contour de la population final comme l'approximation du contour de la masse, l'étape suivante de notre travail est l'extraction des paramètres liés à l'anomalie détectée.

VI.3.3 Extraction des paramètres de la masse

Les masses sont des lésions retrouvées en trois dimensions c'est-à-dire qu'on peut l'observer sur plusieurs incidences. La localisation de la masse est de l'analysée à partir de plusieurs clichés surtout les petites masses qui ne sont pas palpables. La visibilité des contours de la masse évoque la malignité [TA08].

On distingue généralement 5 formes différentes de masses : ronde, ovale, lobulaire, irrégulière ou une distorsion architecturale [HE09].

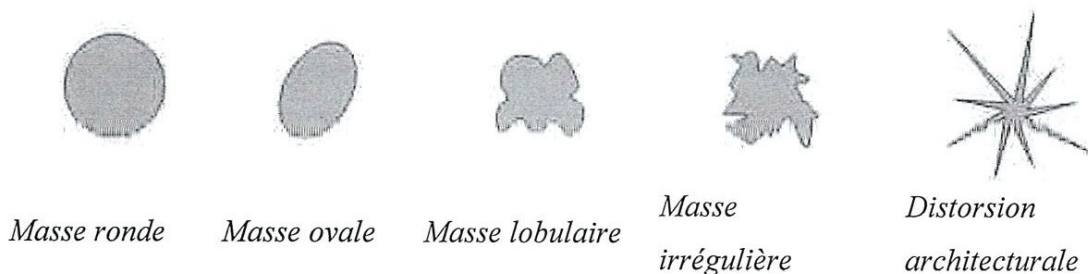


Figure.4.9 Les formes des masses

Les contours de la masse sont très importants à étudier lors de l'étape d'analyse. On distingue généralement 5 types de contours : bien définis, effacés, micro-lobulés, mal définis et avec des spicules. Les contours effacés correspondent à une superposition de tissu adjacent. Les

contours mal définis ou les spicules correspondent à une invasion dans les régions saines [HE09].

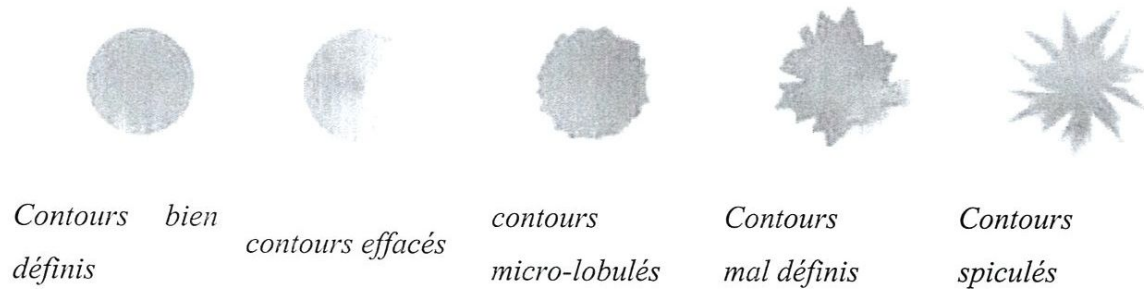


Figure.4.10. Les types des contours des masses.

Les masses sont en général denses (plus blanches) que le tissu mammaire connexe.

VI.3.3.1 Définition des Caractéristique des masses selon BI-RADS

BI-RADS (Breast Imaging System And Data-System) est un system d'aide à la rédaction des compte rendue de plus en plus utilisé dans le monde et bientôt implanté directement sur les matérielles de mammographie et échographie.

Selon BI-RADS les masses sont caractérisé par des descripteurs concernant leurs formes, leurs contour et leurs densité. Pour chaque descripteur, un certain nombre de terme sont définis par un mot et une expression.

Une masse et caractérisé par sa forme (ronde, ovale, lobulé, irrégulière),son type de contour (circonscrit,micro-lobulé,masqué,indistinct,spicululé),sa densité (haute, moyenne, faible).

VI.3.3.2 L'extraction d'informations

Le but de notre travail est d'extraire des informations relatif a la zone englobé par le contour élu comme meilleur dans l'algorithme génétique, pour déterminé les caractéristiques de cette masse il faut extraire les informations suivant :

A- La forme de la masse : la forme de la masse est fortement liée à la courbure et la continuité du contour élu. C'est-à-dire l'énergie interne

La courbure qui a un comportement comme une plaque mince. Il prend une valeur importante lorsque la courbe s'incurve rapidement c'est-à-dire pour l'obtention de coins et les formes complexes, par contre lorsque cette énergie est faible la courbe tendra vers un cercle dans le cas d'un contour fermée [FE09].

B- Type de contours de la masse : le type de contours aussi lié à l'énergie externe du contour élu. Car l'énergie externe prend une valeur importante lorsque la masse a des contours bien définis et circonscrits, par contre on obtient une faible énergie externe dans le cas des contours mal définis ou effacés.

C- La taille de la masse : c'est facilement de calculer la taille « T » de la masse, qui est représenté par la zone englobé par le contour élu.

D- L'intensité de la masse : l'intensité « I » de la masse est égale à la moyenne d'intensités des pixels de la zone englobé par le contour élu.

VI.4 Réalisation

L'interface de notre application se divise en trois fenêtres, chaque fenêtre a une ou plusieurs fonctionnalités, on explique chaque une de ces fenêtres :

VI.4.1 Sélection Du Segment et Définition du Centre de Segment

Cette fenêtre nous offre trois fonctionnalités :

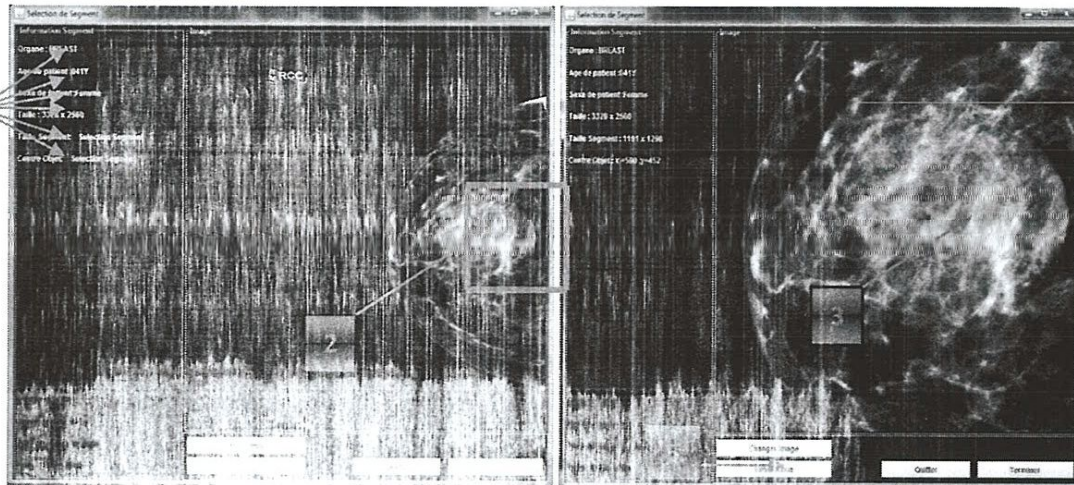


Figure 16 : Sélection du segment et son centre.

- 1) chargement de l'image à l'aide du bouton chargé.
- 2) sélection un segment d'image qu'il nous paré suspect.
- 3) choix du centre du segment sélectionné.
- 4) panneaux d'affichages des informations :

✓ Le nom et l'âge du patient

- ✓ Sexe du patient
- ✓ Taille de l'image originale
- ✓ Taille du segment
- ✓ Coordonné du centre du segment choisi par l'utilisateur

Note : on peut revenir sur nos choix à l'aide du bouton « reselection »(5)

VI.4.2 Prétraitement

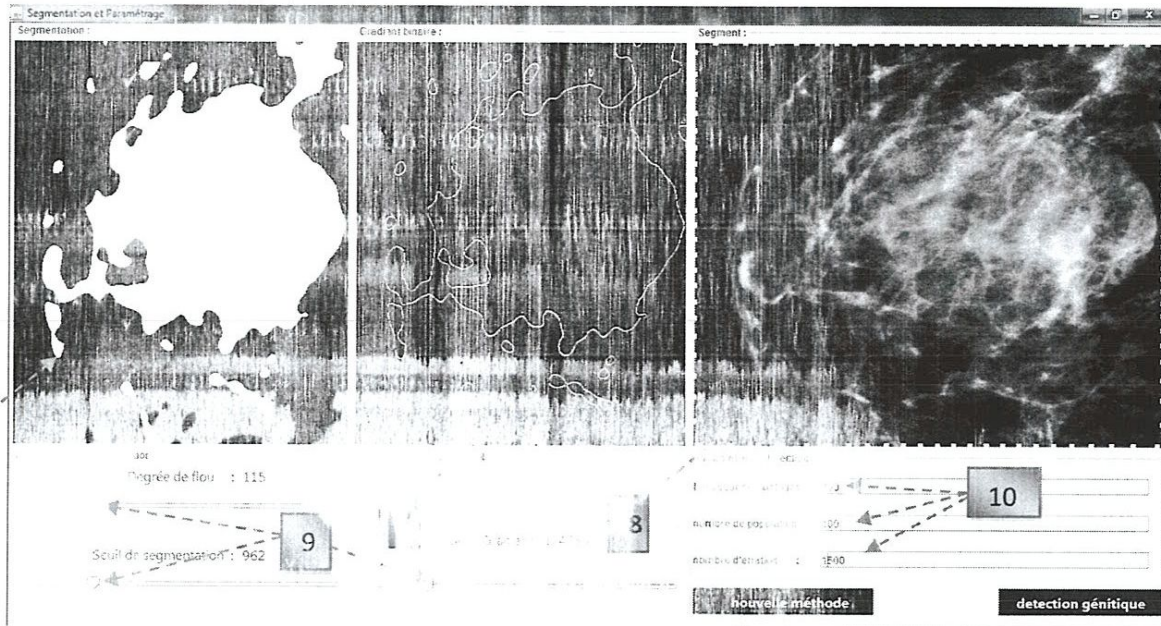


Figure (4.12):selection du segment et son centre.

Cette préparation a pour but de rendre les contours plus apparent

- 6) Résultat de filtrage (moyenneur) + Segmentation par seuillage déjà expliqué dans voir la partie conception
- 7) calcul du gradient de l'image (5).
- 8) Image après prétraitement.
- 9) des « Sliders » pour régler le degré de flou, le seuil de segmentation, et le seuil de gradient.
- 10) des champs de texte pour fixer les paramètres d'entrées de l'algorithme génétique.

VI.4.3 Détection

Après avoir fixé les paramètres de l'algorithme génétique de la fenêtre précédente et à l'aide du bouton détection « génétique » une fenêtre appelé détection apparait :



Figure (4.13): détection du contour.

A) Création de la population initiale, les points jaunes sont celle qui forme le contour.

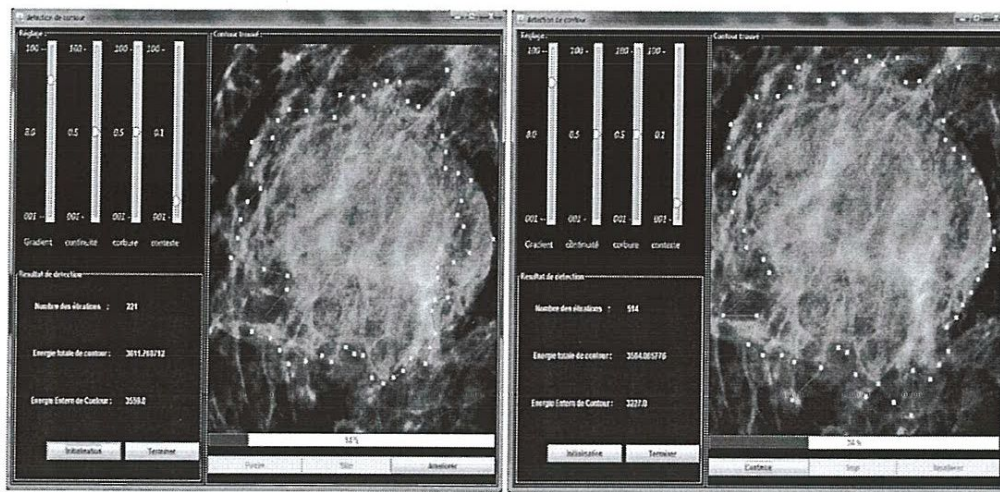
B) Ce bouton sert a lancé la détection.

C) Réglé les coefficients $\alpha, \beta, \gamma, \theta$

VI.4.3.1 Le Déroulement De L'algorithme

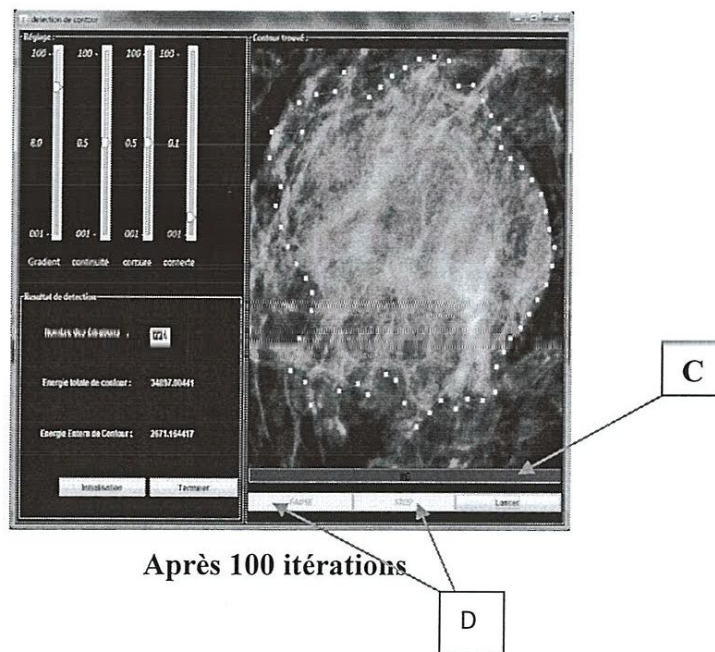
Lorsqu'on clique sur le bouton (b) l'algorithme se débute, la figuré (4.11) montre quelque prise l'ors de son exécution.

On fixe nombre itération =100, onobtient le résultat suivant :



Après 20 itérations

Après 33 itérations



Après 100 itérations

Figure (4.14): le déroulement de l'algorithme génétique.

C) barre de progression pour assuré le « feed-back » de l'interface.

D) si durant le l'exécution on obtient un bon résultat ou le résultat devient stable, on peut arrête le processus a l'aide du bouton stop et le relancé avec continue.

VI.4.3.2 Résultat De L'extraction Des Paramètres

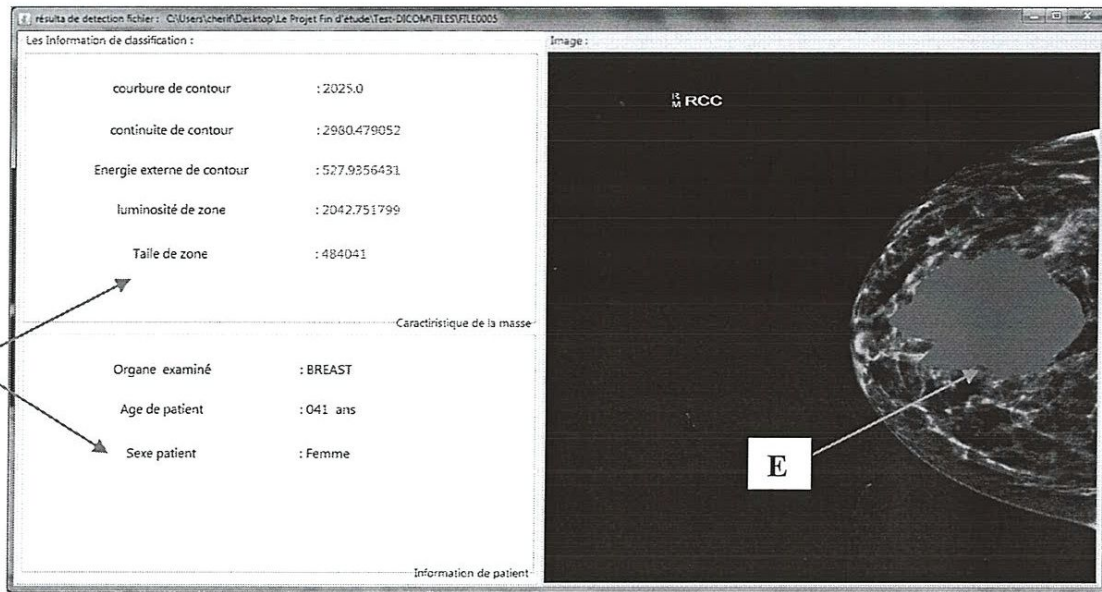


Figure (4.15): Résultat finale.

E) L'anomalie

F) Affichage des informations concernant le patient et ceux de la masse :

Du patient

- Organe examiné
- Sexe et âge patient

La masse

- Courbure
- Continuité
- Energie externe
- Luminosité de la zone
- Taille de la zone

VI.5 étude de cas

Pour arrivé a interprété les résultat,on effectue des tests sur des image simple.

On fixe les paramètres $\alpha = 0,5$ $\beta = 0,5$ $\gamma = 8$ $\theta = 0,1$

Nombre des points qui forme un contour =60 points

Pour un tel paramétrage on obtient les résultats suivants :

VI.5.1 La densité de la masse :

La densité de masse est déterminée par la luminosité de zone englobé par le contour.

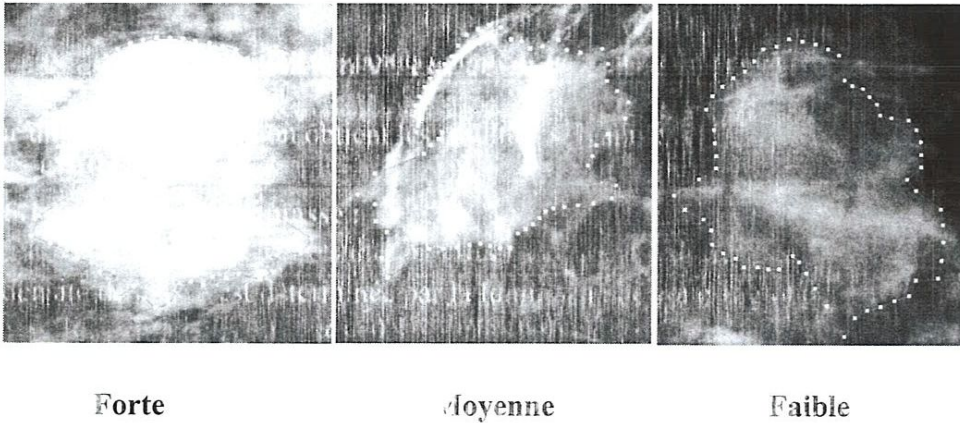


Figure (4.1) La détection sur densité différentes.

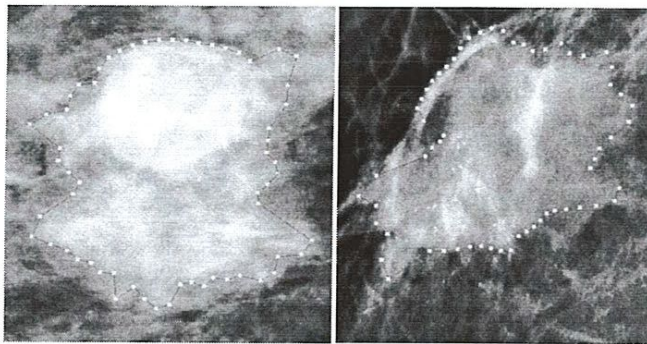
Le tableau (4 .1) présente les différentes valeurs obtenues pour les trois types de densité :

		Les densités de la masse		
		Fort	Moyenne	Faible
Contour	Courbure	2473	2177	1883
	Continuité	3002,5	2632 ,5	2761,7
	Energie externe	684,4	865,5	747,01
La zone	Luminosité	1421,4	1123,5	1810,1
	La taille	87996	97939	36889

Tableau (4.1):Resultat du teste.

VI.5.2 type du contour :

Le type du contour a une relation directe avec la valeur du gradient :



Effacé

Clair

Figure (4.17): contour effacé VS claire.

Avec les mêmes paramètres:

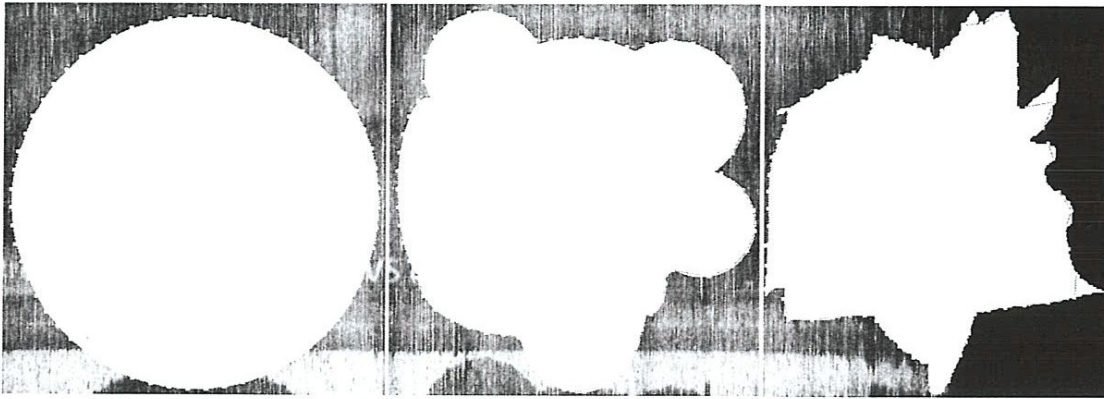
		Type du contour	
		Effacé	Clair
Contour	Courbure	2473	2177
	Continuité	3002,5	2632,5
	Énergie externe	684,4	865,5

Tableau (4.2): les énergies du contour.

Un contour clair se caractérise par un gradient plus élevé. Un contour effacé se caractérise par un gradient faible:

VI.5.3 La Forme de Contours de la Masse

A partir de l'énergie interne de contour qui englobe l'anomalie en détermine sa forme, su les 03 exemples présentés dans la Figure (4.18).



Cercle

Lobulé

Irrégulier

Figure (4.18) : les différentes formes au contour.

le tableau (4.3) montre le résultat des testes .

		Les formes du contour		
		Cercle	Lobulé	Irrégulier
Contour	Courbure	123	1543	2140
	Continuité	3034,84	3262,60	3232,12
	Energie externe	3368,40	1726,65	1682,12
La zone	Imbalance	4067,99	4063,98	4019,48
	La taille	2458074	2143919	2283804

Tableau (4.3): les types de forme.

Si la forme est un cercle l'énergie de courbure et de continuité sont minimale.

Si la forme se déforme en lobulé les deux énergies augmente.et si elle devient irrégulier les énergies seront maximale.

Après une petite étude expérimentale, sur l'énergie de courbure « E_c » on a remarqué que

- $E_c = 123$ la forme est parfaitement un cercle.
- $1543 < E_c < 2140$ la forme est lobulée.
- $E_c > 2140$ la forme devient de plus en plus irrégulière

VI.6 Conclusion & perspective

Dans ce chapitre, nous avons présenté notre travail coté conception et réalisation. Dans la partie conception, nous avons expliqué une méthode de contour actif basé sur les algorithmes génétiques décrite dans le 2eme chapitre en détaillons chaque phases avec des schémas et des pseudos codes; dans la partie réalisation, nous avons présenté l'interface de l'application ainsi que les résultats obtenus à partir d'une image donnée pour aboutir a un résultat fiable.

Dans ce mémoire nous souhaitons réaliser un système d'extraction de paramètres à partir d'une anomalie, le système réalisé est capable d'aider les radiologues pour classifier les masses mammaires, et d'être le point de départ pour la classification automatique.

Les principales perspectives qui apparaissent à l'issue de ce travail concernent la classification des anomalies à travers la conception d'un système de classification, ainsi que l'amélioration du module de paramétrage des contours actifs génétiques pour le rendre automatique, aussi en faisant appel à des systèmes de segmentation existants. Afin d'améliorer les résultats d'extraction des caractéristiques.

CONCLUSION GENERALE

La détection de contours joue un rôle important lorsqu'on cherche à résoudre un problème d'analyse d'images, les méthodes des contours actifs abordées dans le premier chapitre prennent une place tout à fait particulière parmi les techniques de détection de contours, où nous avons présenté les implémentations les plus connues. Il s'agit dans tous les types de Snakes de minimiser une somme d'énergies.

Le cancer du sein est le cancer le plus fréquent chez la femme, en Algérie le nombre de décès annuels dus à cette maladie est évalué à 3500 selon les statistiques de l'institut national de santé publique Algérienne. En outre, on recense 7000 nouveaux cas de maladie par an.

Le problème que nous venons de résoudre est un problème complexe à cause du caractère difficile des mammographies en terme de netteté de contours et de degré de contraste, où la qualité des images joue un rôle important dans le processus de détection des lésions dans les mammographies. Les recherches dans ce domaine sont nombreuses mais les résultats obtenus aujourd'hui sont encore assez peu avancés.

L'architecture générale que nous avons proposée pour la résolution de ce problème est composée de trois étapes principales, dans la première étape nous avons utilisé quelques étapes de prétraitement, la deuxième étape de notre système consiste à détecter les contours de la masse avec les contours actifs génétiques, cette méthode consiste à utiliser la puissance exploratoire des algorithmes génétiques. Cette implémentation utilise un algorithme génétique pour faire évoluer une population de snakes vers un contour qui présente une énergie minimum après un nombre d'itération, ceci est considéré comme le contour approximatif de l'anomalie. La troisième étape consiste ensuite à analyser le contour élu de l'étape précédente pour l'extraction de caractéristiques liées à la masse comme la forme, la taille, le type de contours et l'intensité de l'anomalie.

Nous avons testé notre algorithme sur des mammographies numériques format DICOM codé sur 16 bits, et sur différents types d'anomalies, les résultats que nous avons obtenus sont encourageants et permettent d'espérer un aboutissement heureux. Une interaction de l'utilisateur est nécessaire pour déterminer le paramétrage des contours actifs génétiques, pour cela il faut déterminer une méthode de paramétrage automatique, nous devons aussi réaliser un système de classification au lieu de donner des valeurs numériques, ces constatations ouvrent donc des perspectives à de nouvelles recherches.

-----Bibliographie-----

[BOU09] Boudermine.R,« Détection des cavités cardiaques en utilisant les contours actifs »
Mémoire de fin d'études Pour l'obtention du diplôme d'ingénieur d'état en informatique.
École National Supérieure d'Informatique (E.S.I), Alger, 2008/2009.

[BL99] Ballerini, L,«Genetic Snakes for Medical Images Segmentation. Lectures Notes in
Computer Science»,Livre sur la segmentation des images médicales. vol. 1596, p. 59-73.
1999.

[BT93] Bäck, T,« Optimal Mutation Rates in Genetic Search Fifth»,Conférence inter
nationale sur les algorithmes génétique,San Mateo, CA, USA, 1993.

[CF88] Christine. F,«Introduction aux techniques de traitement et d'analyse d'images par
Christine FERNANDEZ-MALOIGNE Professeur - Signaux et Images Multi sources et
Multi composantes Laboratoire IRCOM-SIC, UMR 6615 CNRS»,Université de Poitiers,
1988.

[COU05] Antoine.C, « La méthode des contours actifs en traitement des images »,Mémoire pour
l'examen probatoire en Calcul Scientifique, Conservatoire National des Arts et Métiers, Paris,
Le 02 février 2005.

[DM11] Dridi S-E & Messiod B,« Les modèles Déformables pour la Détection de
Contours», Mémoire de fin d'études pour l'obtention du diplôme licence en informatique,
Université de Guelma, Faculté des Mathématiques, d'Informatique et des Sciences de la
matière, Juin 2011.

[DJ96] Deccaloc, J. L, «L'ordinateur génétique »,Livre sur les algorithmes génétique,
Edition : Hermès, Paris 1996.

[FE09] Fekir A & Benarrane N & Taleb-Ahmed A, « Détection et suivi d'objets dans une séquence d'images par contours actifs », Conférence Internationale sur l'Informatique et ses Applications, Vol.2, Algérie, 2009

[GW97] R.C. Gonzales & P. Wintz, « Digital Image Processing », Livre sur le traitement d'image, Edition: Addison Wesley 1997.

[JM08] J-m Munguakonkwa Biringanine, « La liaison automatique des plusieurs images perçues sur un scanner », Mémoire de fin d'études pour l'obtention du diplôme licence en pédagogie; Option : Informatique de Gestion, (ISP) Institut Supérieur Pédagogique de Bukavu (ISP) 2008 .

[JN05] J. Alliot & N. Durand, « Méthode des Algorithmes génétiques », Article sur les AGs publié le : 14 mars 2005 .

[KH08] Kaddour.B & Hamtouche.A, « Contours actifs paramétriques pour la segmentation d'images et vidéos », Mémoire de fin d'études pour l'obtention du diplôme d'ingénieur d'état informatique, Institut National de formation en Informatique (I.N.I), Alger, 2007/2008.

[Ho75] HOLLAND, J. H. - Adaptation in Natural and Artificial Systems. University of Michigan Press, 1975, Livre sur les contour actif génétique (Second édition : MIT Press, 1992).

[HK10] Hocini.L & Kori yahia S.D, « Détection de mouvement dans une séquence vidéo par une méthode multimodale », Mémoire de fin d'études pour l'obtention du diplôme d'ingénieur d'état en informatique, Institut National de formation en Informatique (I.N.I), Alger, 2009/2010.

[HE09] Heron J.F, « La Mammographie; Cancérologie générale », www.oncoprof.net, 2009.

[KC99] K.Chakib, « Généralités Sur Le Traitement D'images », Livre sur le traitement numérique des images, Edition: Addison Wesley, 1999.

[LA91] Lissouaoui.N&Hamami .L & Nouali.N, « Les algorithmes génétiques application a la segmentation des images = Genetic algorithm : Application to image segmentation», Revue : RIST Vol 14 N°02 Année 2004,Centre de recherche sur l'information scientifique et technique, Alger, ALGERIE 1991 .

[LL05] L. Lévy, M. Suissa, J. Bokobsa, H. Tristant, J.-F. Chiche, B. Martin, G. Teman, «Breast Imaging Reporting System and Data System».Gynécologie Obstétrique & Fertilité 33 (2005) 338–347.

[MA87] Marion. A,« Introduction aux techniques de traitement d'images »,Livre sur le traitement d'images, Edition Eyrolles, 1987.

[MZ92] Michalewicz, Z,«Genetic Algorithms + Data Structures = Evolution Programs»,Livre sur les algorithmes génétique,Edition : Springer-Verlag - Berlin-Heidelberg, 1992.

[TA08] Tardivon A « Les images détectées », le dépistage du cancer du sein: un enjeu de santé publique, Vol.2, springer, pp.115-129, 2008.

[RR09] Rabehanta Nirina.N & Rasoloarison Tiana.M,« Approches de détection des contours : implémentation et comparaison » Mémoire de fin d'études pour l'obtention du diplôme licence en informatique,Université de Guelma, faculté des sciences et de l'ingénierie, Juin 2009.

[ROU03] J-J.Rousselle,« Les contours actifs, une méthode de segmentation Application à l'imagerie médicale », Thèse pour obtenir le grade de docteur de l'université de tours, Université François Rabelais de Tours, Le 09 juillet 2003.

[PS08] P.Schwartz,«Article sur Méthode des contours actifs», Publier le 5 mars 2008 .

[ST06] Samamba.T,« Reconnaissance des formes comme outil d'aide aux traitements d'image. Cas des empreintes digitales»,Mémoire de fin d'études pour l'obtention du diplôme master , Option :informatique, (ISP) Institut Supérieur Pédagogique de Bukavu,2005-2006.

[SR04] Souquet.A & Radet .F,«Algorithmes Génétiques»,Thèse de fin d'année Tutorat de Mr Philippe Audebaud, Université Paris-Sud, France, Soutenu le 21/06/2004.

-----Site Web-----

[1] « détection de contours »

[Http://docs.google.com/](http://docs.google.com/): page consultée le : 04/02/2011

[2] « Filtrage de Canny »

http://www.optique-ingenieur.org/fr/cours/OPI_fr_M04_C05/co/Contenu_02.html.

[3] «Algorithmes Génétiques»

http://eviewbox.sourceforge.net/JFR98/Algorithme_génétique,consulté le : 04/02/2012.

[4] «New Genetic Local Search Operators for the Traveling Salesman Problem»

[http:// http://citeseer.nj.nec.com/freisleben96new.html](http://citeseer.nj.nec.com/freisleben96new.html) consultée le : 07/05/2012.

[5] «la norme DI-COM »

<http://eviewbox.sourceforge.net/JFR98/intro.html> consultée le : 24/05/2012.