

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

UNIVERSITE MOULOU D MAMMERRI DE TIZI-OUZOU



FACULTE DU GENIE ELECTRIQUE ET D'INFORMATIQUE

DEPARTEMENT D'INFORMATIQUE

Mémoire de Fin d'Etudes de MASTER ACADEMIQUE

Domaine : **Mathématiques et Informatique**

Filière : **Informatique**

Spécialité : **Conduite des Projets Informatiques**

Thème :

Détection d'objet basé sur l'estimation de la saillance

Présenté par :

M^{elle} Katia DJAILEB
M^{elle} Wahiba OUALLI

Encadré par :

Mr FILALI Idir

Promotion 2017/2018

Remerciements

D'abord, nous remercions le bon DIEU de nous avoir donné santé et courage pour réaliser ce modeste travail.

Nous tenons à exprimer notre profonde gratitude à notre promoteur Mr FILALI Idir, pour nous avoir encadrés et guidés et pour son aide précieux, ses conseils durant toute la durée de notre travail.

Nos vifs remerciements vont également aux membres de jury pour l'honneur qu'ils nous font en acceptant de juger notre mémoire de fin d'études Et de l'enrichir par leurs propositions.

Nos sincères sentiments vont à nos familles qui ont sacrifié jusqu'aujourd'hui et leurs encouragements incessants tout le long de notre parcours.

Katia , Wahiba

Dédicace

À tous ceux qui me sont chers et proches ;

À mes parents et ma grand-mère :

À ma sœur et mes frères ;

À mes amies ;

*À ma binôme **Wahiba** et toute sa famille ;*

Enfin à tous ceux que j'aime,

Je dédie ce travail.

Katia

Dédicace

*A mes très chers parents « **Rabah & Aldjia** » qui ont toujours été là pour moi, qui m'ont donné un magnifique modèle de labeur et de persévérance, pour leur attention, sacrifice et soutien tout au long de mes études ;*

*A mes très adorables frères : **Hamid & Mahdi** ;*

*A mes très chères sœurs : **Samira & Razika** et son époux **Mustapha** ;*

*A ma binôme **Katia** et toute sa famille ;*

*A mes chères amies : **Djamila & Houria & Katia & Dyhia & Radia** ;*

Enfin à tous ceux qui me connaissent de près ou de loin.

Wahiba

| | |
|---|----|
| INTRODUCTION GENERALE..... | 1 |
| Chapitre 1 : GENERALITES SUR LES IMAGES..... | 4 |
| I Introduction :..... | 5 |
| II Représentation d’une image :..... | 5 |
| II.1 Notion d’image :..... | 5 |
| II.2 Espace de couleur :..... | 5 |
| III Représentation d’une image numérique :..... | 7 |
| III.1 Les images matricielles (ou bitmap):..... | 7 |
| III.2 Les images vectorielles :..... | 7 |
| IV Caractéristique d’une image numérique :..... | 8 |
| IV.1 Pixel :..... | 8 |
| IV.2 Bruit :..... | 8 |
| IV.3 Dimension :..... | 9 |
| IV.4 Le contraste :..... | 9 |
| IV.5 Résolution :..... | 9 |
| IV.6 Extension :..... | 9 |
| IV.7 La luminance :..... | 9 |
| V Conclusion :..... | 10 |
| Chapitres 2 : LES PRETRAITEMENTS DE L’IMAGE NUMERIQUE | 11 |
| I Introduction :..... | 12 |
| II Les prétraitements :..... | 12 |
| II.1 Modifications de l’histogramme :..... | 12 |
| III Les filtrages :..... | 16 |
| III.1 Filtrages linéaires :..... | 17 |
| III.2 Filtrages non linéaires :..... | 21 |
| III.3 Filtrage linéaire itérés :..... | 23 |

| | | |
|-----------------------------------|--|----|
| IV | Les opérateurs morpho mathématiques (ou morphologiques) : | 24 |
| IV.1 | La dilatation : | 25 |
| IV.2 | L'érosion : | 25 |
| V | Correction de l'illumination : | 26 |
| V.1 | Correction prospective : | 26 |
| V.2 | Correction rétrospective : | 29 |
| VI | Conclusion : | 30 |
| Chapitre 3 : SAILLANCE D'OBJET | | 31 |
| I | Introduction : | 32 |
| II | Définition de la saillance : | 32 |
| III | La détection d'objet saillant : | 32 |
| IV | Les facteurs de la saillance visuelle : | 33 |
| IV.1 | Facteurs physiques : | 33 |
| IV.2 | Facteurs cognitifs : | 34 |
| V | Les méthodes de la saillance : | 36 |
| V.1 | Méthodes basées sur l'estimation du contraste local : | 36 |
| V.2 | Méthodes basées sur l'estimation du contraste global : | 36 |
| V.3 | Méthodes hybrides : | 40 |
| VI | Conclusion : | 44 |
| Chapitre 4 : SEGMENTATION D'OBJET | | 45 |
| I | Introduction : | 46 |
| II | Définition de la segmentation : | 46 |
| III | Méthodes de segmentation : | 46 |
| III.1 | La méthode de segmentation par seuillage (la binarisation) : | 46 |
| III.2 | La méthode de segmentation par croissance de régions : | 48 |
| III.3 | La méthode de segmentation par les contours : | 49 |

| | | |
|--|--|----|
| III.4 | Les méthodes de décomposition/ fusion (split and merge) :..... | 50 |
| III.5 | Segmentation avec Graph cut (coupe de graph) :..... | 52 |
| IV | Conclusion :..... | 54 |
| Chapitre 5 : Choix de la méthode & mise en œuvre | | 55 |
| I | Introduction :..... | 56 |
| II | Environnement du travail :..... | 56 |
| III | La détection de la saillance avec Chaîne de Markov absorbante :..... | 57 |
| IV | La segmentation avec « graph_cut » :..... | 61 |
| V | Evaluation de la méthode proposée :..... | 65 |
| VI | Résultats de segmentation par les trois méthodes :..... | 66 |
| VII | Conclusion :..... | 66 |
| CONCLUSION GENERALE | | 67 |
| Références bibliographique..... | | 69 |

| | |
|---|----|
| Figure 1: Trois classes d'images numériques | 6 |
| Figure 2: Exemple des images numérique [5]..... | 7 |
| Figure 3 : Exemple d'un pixel (groupe de pixel formant la lettre A) [6]..... | 8 |
| Figure 4: Exemple d'une image et son histogramme..... | 13 |
| Figure 5: Interprétation d'un histogramme | 13 |
| Figure 6: Etalement d'histogramme | 14 |
| Figure 7: L'étalement d'histogramme..... | 15 |
| Figure 8: Egalisation d'histogramme | 15 |
| Figure 9: L'égalisation d'histogramme | 16 |
| Figure 10: Illustration du lissage par filtre moyenneur avec différents valeur de voisinage W | 17 |
| Figure 11: Représentation graphique d'une fonction gaussienne pour $\sigma = 1$ | 18 |
| Figure 12: Illustration du lissage par filtre gaussien avec différents valeur de voisinage W | 19 |
| Figure 13: Illustration de lissage par filtre gaussien et filtre moyenneur avec la même valeur de voisinage $W = 5 \times 5$ | 19 |
| Figure 14: Exemple d'un filtre Sobol..... | 20 |
| Figure 15: Exemple d'un filtre laplacien..... | 21 |
| Figure 16: Exemple de masque Laplacien 8-connexe appliqué a une image..... | 21 |
| Figure 17 : Illustration du principe du filtre médiane avec une matrice de taille $W=3 \times 3$ | 22 |
| Figure 18: Illustration du lissage par filtre médian avec différents valeur de voisinage W | 22 |
| Figure 19: Application du filtre gaussien 3×3 avec $\sigma = 1$:..... | 23 |
| Figure 20: Exemples d'élément structurant | 24 |
| Figure 21: Dilatation d'image | 25 |
| Figure 22: Erosion d'image..... | 26 |
| Figure 23: Correction à l'aide des images noire et blanche..... | 27 |
| Figure 24 : Correction à l'aide de l'image blanche seule | 28 |
| Figure 25 : Correction à l'aide de l'image noire seule | 29 |
| Figure 26 : Correction rétrospective à base de filtre passe-bas..... | 30 |
| Figure 27: Exemple illustrant les différentes résolutions du graphe proposé dans [17]. | 43 |
| Figure 28 : Segmentation d'images par seuillage | 48 |
| Figure 29 : segmentation par croissance de régions..... | 49 |
| Figure 30 : segmentation par les contours (https://perso.telecom-paristech.fr) | 50 |

| | |
|--|----|
| Figure 31 : Split de l'image I [46]. | 51 |
| Figure 32: Merge de l'image I [46]. | 51 |
| Figure 33: Exemple d'un graphe [47]. | 52 |
| Figure 34: Exemple d'une coupe de graphe [47]. | 53 |
| Figure 35: Exemple d'un graph $S-T$ [47]. | 53 |
| Figure 36: La page d'accueil de MATLAB R2014a. | 57 |
| Figure 37 : fenêtre principal de matlab | 57 |
| Figure 38 : les images générées avec « chaine de markov » | 61 |
| Figure 39 : Exemple d'un graphe et les liens pondérées, avec p, q, r des | 62 |
| Figure 40 : image segmenter par graph-cut. | 64 |
| Figure 41 : image segmentée par « otsu » et « multithresh » | 66 |

RVB : Rouge, Vert, Bleu.

CIE : Commission Internationale de l'Eclairage.

YUV : Y est la luminance, U et V contiennent l'information sur la couleur (forment la chrominance de la couleur).

PIXEL: Picture Elements.

BMP: Bit map.

JPG: (JPEG) Joint Photographic Experts Group.

GIE : groupement d'intérêt économique

PNG : Portable Network Graphics.

SVG : Le Scalable Vector Graphics (en français « graphique vectoriel adaptable »), est un format de données conçu pour décrire des ensembles de graphiques vectoriels et basé sur XML

DXF : Drawing eXchange Format.

DWG : DraWinG est le format natif des fichiers de dessins AutoCAD

MIN: Minimum.

MAX: Maximum.

CIE LAB: (CIE L^*a^*b) est un espace de couleur particulièrement utilisé pour la caractérisation des couleurs de surface.

Graph cut: Coupe de graph.

MATLAB: Matrix Laboratory.

Résumé :

De nos jours, le traitement d'image a connu d'éventuelles motivations qui touchent plusieurs domaines de la vie quotidienne.

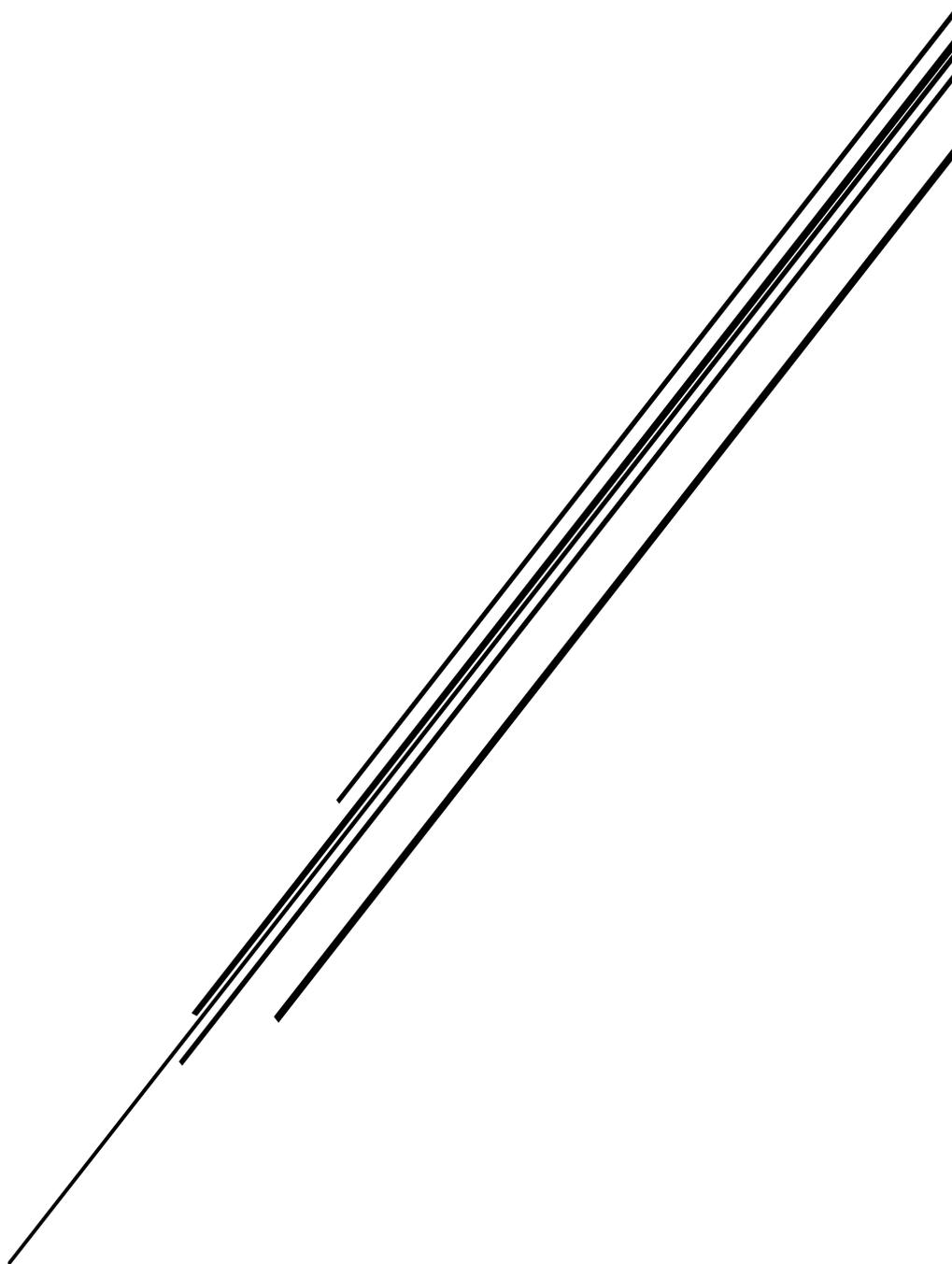
Dans ce contexte, nous avons abordé une approche indispensable dans le traitement d'image qui est détection d'objet basées sur l'estimation de la saillance.

Nous nous intéressant dans ce travail à l'une des approches de la segmentation basée sur : au premier lieu détecter l'objet saillant avec « chaine de markov », ensuite nous proposons une approche qui base sur la théorie des graphes qui considère l'image à segmenter comme un graphe « Graph-cut»,

Les résultats obtenus sont très satisfaisants et montrent une bonne détection d'objet saillant dans une image.

Mots-clés : saillance « chaine de markov », segmentation « Graph-cut »

INTRODUCTION GENERALE



Au cours de développement de la science, on ne peut pas nier le rôle de l'image numérique, un support important dans différents domaines tel que l'imagerie médicale, la photographie, l'image satellitaire et les documents numériques.

Le traitement d'image est un domaine qui traite les images numériques afin d'éliminer le bruit, les corriger et améliorer leurs qualités. Nous avons ainsi plusieurs techniques de traitement qui sont les filtres, les modifications d'histogramme et les opérations morpho mathématiques...

Les êtres humains possèdent la faculté d'identifier rapidement des objets saillants dans des scènes visuelles complexes avec une précision remarquable. La notion de saillance contient une idée centrale, mais le détail varie selon le domaine d'application. Un objet est saillant s'il est facilement remarqué. A partir de cette idée intuitive, plusieurs définitions peuvent être proposées, selon le critère qui fait que l'objet va être remarqué.

La segmentation d'image intervient dans la plupart des applications de traitement d'image (reconnaissance d'objet, détection d'un objet dans une image...etc.), elle permet de trouver les régions homogènes d'une image ou les contours qui séparent ses régions hétérogènes et la segmentation d'un objet. Il existe une multitude de techniques de segmentation d'image, segmentation par région, par contour et par seuillage.

Dans notre travail, nous allons nous baser sur la saillance des objets dans une image et la segmentation de cet objet, ainsi notre objectif est de détecter les objets saillants dans une image

Afin d'arriver à cet objectif, et de donner une démarche compréhensible, nous avons structuré le présent mémoire de la manière suivante :

Chapitre I : Intitulé « Généralité sur les images » donne un aperçu global sur la notion des images et leurs différents types et caractéristiques.

Chapitre II : Intitulé « les prétraitements de l'image numérique » dans ce chapitre nous parlons d'une manière générale sur le traitement de l'image, qu'on peut appliquer sur ces images. Nous avons présenté les différents filtres, la modification de l'histogramme, les opérateurs morpho mathématiques ainsi la correction de l'illumination.

Chapitre III : Intitulé « Saillance d'objet » donne un aperçu général sur la notion de saillance, les objets saillants, et les différents facteurs et méthodes existant.

Chapitre IV : Intitulé « Segmentation des objets » présente la notion de segmentation d'objet, les différentes méthodes utilisées pour atteindre les objectifs voulus.

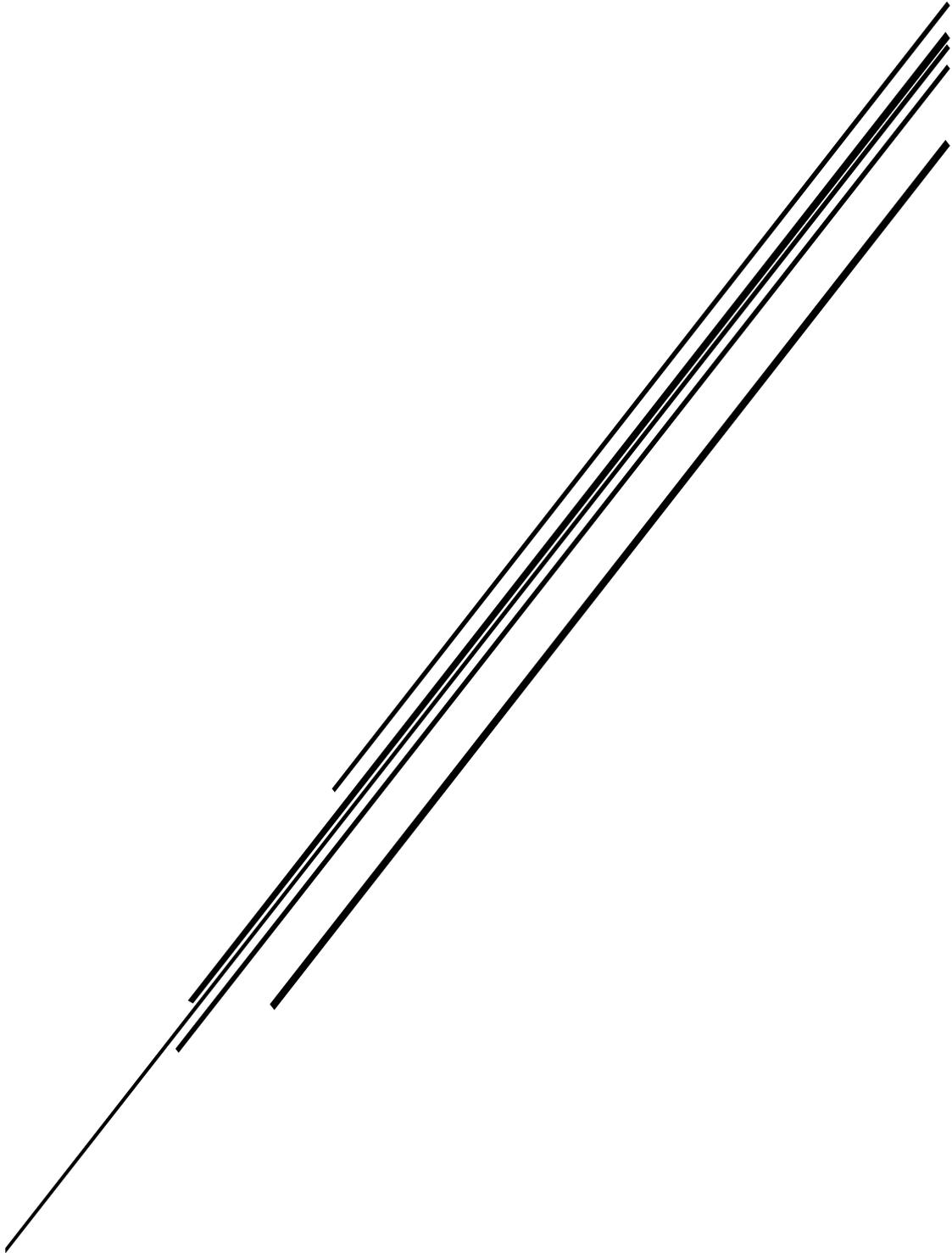
Chapitre V : Intitulé « Choix de la méthode et mise en œuvre » contient la description des étapes suivies et les méthodes utilisées pour détecter l'objet saillant et par suite le segmenter, et à la fin nous avons évalué les méthodes choisies.

Problématique : comment détecter un objet saillant ? Quelle est la meilleure méthode pour le faire ? Est-il possible de segmenter un objet dans une image ? Quelle est la méthode qui donne un bon résultat ?

Enfin, la conclusion générale résumera nos contributions et donnera quelques perspectives sur les travaux futurs

Chapitre 1 : GENERALITES

SUR LES IMAGES



I Introduction :

L'image occupe une place très importante dans le monde moderne, elle intervient fréquemment dans notre vie quotidienne (téléphone, télévision, magazines...), professionnelle (vision industrielle, imagerie médicale, robotique ...), il est ainsi intéressant de comprendre la structure de cette image sous sa forme numérique ainsi que ses différentes caractéristiques.

II Représentation d'une image :

II.1 Notion d'image :

L'image est une reproduction significative d'ensemble d'information structuré caractérisant un univers composé d'entités : objets figurant dans une scène réelle, la sculpture, le dessin, la photographie, le film, ...etc.

C'est aussi un ensemble structuré d'informations qui a une signification pour l'œil humain, après l'affichage sur l'écran.

L'image peut être écrite 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 la fonction d'intensité lumineuse et de couleur. Alors, sous cet aspect, l'image est inexploitable par la machine, ce qui nécessite sa numérisation [1].

II.2 Espace de couleur :

On peut définir trois classes d'espace de couleur :

- L'image en niveaux de gris ;
- L'image binaire (noir et blanc) ;
- L'image en couleur (RVB), (CIE), (YUV).

II.2.1 Image en niveau de gris :

Le niveau de gris dans une image numérique représente l'intensité lumineuse d'un pixel lorsque ses composantes de couleur sont identiques en intensité lumineuse.

Pour convertir une image couleur en niveau de gris pourrait être de calculer les moyennes des trois composantes (RVB) [2].

La conversion d'une image couleur en une image en niveau de gris se fait suite à l'application de la fonction suivante :

$$G(i, j) = \frac{[R(i, j) + V(i, j) + B(i, j)]}{3} \quad \mathbf{1.1}$$

Avec :

$G(i, j)$: Niveau de gris du pixel situé à la ligne i et à la colonne j .

$R(i, j)$: Intensité de la couleur rouge du pixel (i, j) .

$V(i, j)$: Intensité de la couleur verte du pixel (i, j) .

$B(i, j)$: Intensité de la couleur bleu du pixel (i, j) .

II.2.2 Image binaire :

Binariser une image c'est convertir une image (signal) analogique non idéale délivré par un capteur en une image binaire idéale.

La binarisation d'image est la méthode la plus simple de segmentation d'image, elle consiste à réduire le nombre de niveau de gris d'une image à deux. Cette opération est principalement basée sur le seuillage. Le choix d'un seuil S à partir de l'histogramme permet la division de l'ensemble des niveaux de gris de l'image en deux classes, on puisse étiqueter « blanc » tous les pixels de niveau de gris supérieur ou égal à ce seuil et « noir » tous les pixels de niveau de gris inférieur à ce même seuil.

Dans la pratique, cette situation idéale ne se rencontre que très rarement [3].

II.2.3 Image en couleur (modèle RVB) :

Une image couleur est en réalité composée de trois images indépendantes, afin de représenter le rouge, le vert, et le bleu. Chacune de ces trois images s'appelle un canal.

Chaque pixel de l'image couleur contient ainsi trois nombres (R, V, B) , chacun étant un nombre entier entre 0 et 255. Si le pixel est égal à $(R, V, B) = (255, 0, 0)$, il ne contient que de l'information rouge, et est affiché comme du rouge. De façon similaire, les pixels valant $(0, 255, 0)$ et $(0, 0, 255)$ sont respectivement affichés vert, et bleu.



Figure 1: Trois classes d'images numériques

Image en couleur, image en niveaux de gris, image binaire

III Représentation d'une image numérique :

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

Une image numérique est un signal numérique contient un nombre fini de points. Ces point sont appelés pixels (contraction des mots anglais "Picture Element", c'est à dire élément d'image) qui représentent chacun une portion de l'image.

L'image numérique se représente sous la forme d'un tableau (matrice) à deux dimensions contenant des valeurs entières pour les images en niveaux de gris. Une image ne comportant que 2 niveaux de gris : 0 (noir) et 1 (blanc) est appelée une image binaire, ou des triplets de valeurs entières pour les images couleurs (Rouge, Vert, Bleu).

On distingue deux types d'images numériques: les images matricielles (ou image bitmaps) et les images vectorielles [4].

III.1 Les images matricielles (ou bitmap):

Une image matricielle est composée comme son nom l'indique d'une matrice de points à plusieurs dimensions (un tableau a deux dimensions), ces points sont t'appelé des pixels.

L'image matricielle utilisée dans le domaine du traitement et de l'analyse d'images. Il existe de nombreux formats de fichiers d'extension : **.BMP**, **.JPG**, **.GIF**, **.PNG** sont des fichiers d'images bitmap.

III.2 Les images vectorielles :

Une image vectorielle en informatique, est une image numérique composée d'objets géométriques individuels (segments de droite, polygones, arcs de cercle, etc.) repérés par leurs coordonnées et comportant différentes attributs (de forme, de position, de bordure, de couleur, etc.).

L'image vectorielle utilisée principalement dans le monde du graphisme et de la conception assistée par ordinateur.

Il existe de nombreux formats de fichiers vectoriels, parmi lesquels on peut citer : **.SVG**, **.DXF**, **.DWG**



Figure 2: Exemple des images numérique [5].

A) Image matricielle (bitmap). B) image vectorielle

IV Caractéristique d'une image numérique :

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

IV.1 Pixel :

Une image est constituée d'un ensemble de points appelés pixels, contraction de l'expression anglaise « picture elements » éléments d'image. Le pixel représente ainsi le plus petit élément constitutif d'une image numérique, c'est une entité calculable qui peut recevoir une structure et une quantification. L'ensemble de ces pixels est contenu dans un tableau à deux dimensions constituant l'image.

La quantité d'information que véhicule chaque pixel donne des nuances entre les images monochromes et les images couleurs. Dans les images monochromes, chaque pixel est codé sur un octet. Et 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), bleu(B) [1].

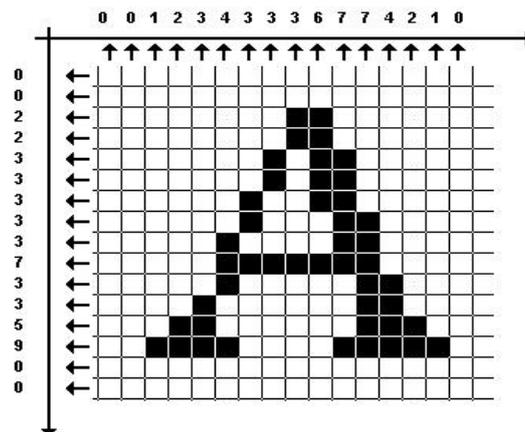


Figure 3 : Exemple d'un pixel (groupe de pixel formant la lettre A) [6].

IV.2 Bruit :

Le bruit est très souvent présent sur les images de documents, car il peut apparaître à différents endroits de la chaîne de numérisation : à l'impression, pendant la vie du document et à la numérisation.

Des traitements classiques tels que les filtres médians sont utilisés pour filtrer des pixels isolés sur l'image. Les opérateurs de morphologie mathématique sont également très souvent utilisés pour corriger de légers défauts sur les images comme "recoller" des morceaux de caractères [7].

IV.3 Dimension :

La dimension est la taille de l'image, se présentée sous forme de matrice dont les éléments sont des valeurs numériques représentative des intensités lumineuses (pixel), le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre totale de pixels dans une image [7].

IV.4 Le 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.

Si L1 et L2 sont les degrés de luminosité respectivement de deux zones voisines A1 et A2 d'une image, le contraste C est défini par le rapport suivant [8] :

$$C = \frac{L1 - L2}{L1 + L2} \quad 1.2$$

IV.5 Résolution :

Le mot résolution est la taille d'une image (nombre de pixels), utilisé pour désigner le nombre totale de pixels affichables horizontalement ou verticalement, où ce nombre est plus grand, la résolution est meilleure [2].

IV.6 Extension :

L'extension d'un fichier, c'est le point, et les trois ou quatre lettres qui suivent le nom d'un fichier. Grâce à ces quelques lettres, vous pouvez connaître le type du fichier (application, documents texte, page web, image, etc). Il suffit d'un clic droit sur le fichier et de lire les propriétés pour trouver l'extension de ce fichier [9].

- Exemple d'extension :
 - Word.exe = application
 - Document.txt = document texte
 - Index.htm = page web
 - Image.jpg (.gif) = image

IV.7 La luminance :

La luminance est tout simplement la sensation visuelle qu'on a de la lumière. De toutes les grandeurs qui caractérisent la lumière, la luminance est celle qui nous concerne le plus,

lorsque la vision est gênée dans une forte pénombre, ou lorsqu'un phare nous éblouit, c'est la luminance qui est en cause.

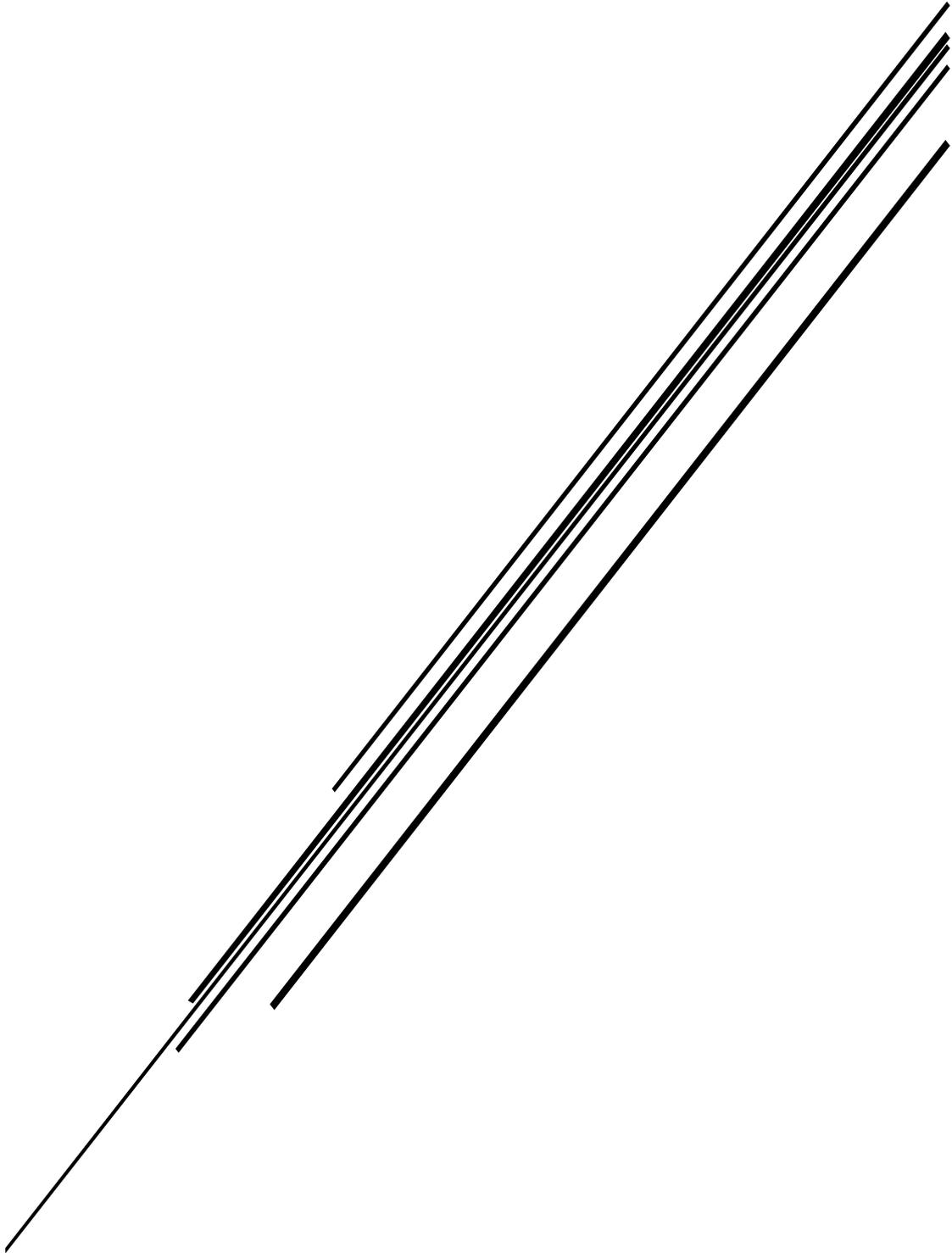
Il y a plusieurs facteurs qui influencent le niveau de luminance. En premier bien sûr, l'intensité lumineuse de la source de lumière. Ensuite lorsque la luminance est renvoyée par une surface éclairée. Et enfin, la coloration de la source lumineuse influence aussi le niveau de luminance [10].

V Conclusion :

Dans ce premier chapitre, nous avons présenté l'image numérique d'une manière générale, et ces différentes représentations au point de vue de la machine, ainsi que les caractéristiques que nous pouvons retirer de ces images.

Le chapitre suivant sera consacré au quelque traitement que nous pouvons appliquer à une image numérique.

**Chapitres 2 : LES
PRETRAITEMENTS DE
L'IMAGE NUMERIQUE**



I Introduction :

Le prétraitement d'image est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années. On désigne par prétraitement d'image numérique l'ensemble des techniques permettant de modifier une image numérique afin d'améliorer ou d'en extraire des informations.

II Les prétraitements :

Les prétraitements d'images est une discipline de l'informatique qui étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information.

Ces prétraitements consistent d'une part à modifier l'apparence d'une image pour qu'un observateur puisse plus facilement en extraire des informations, et d'autre part à supprimer l'information inutile (le bruit) des images pour renforcer l'information utile en vue d'un traitement ultérieur.

Les principaux prétraitements qu'on applique sur les images de documents sont :

- Les modifications d'histogramme ;
- Les filtrages ;
- Les opérations morpho mathématiques ;
- correction de l'illumination ;

II.1 Modifications de l'histogramme :

On cherche à modifier l'image en lui appliquant une transformation ponctuelle d'intensité. Du fait de leur caractère ponctuel (les pixels sont traités individuellement), les méthodes de transformation d'histogramme n'affectent pas la forme des régions mais modifient uniquement l'apparence visuelle de l'image.

Parmi plusieurs types de modification d'histogramme nous citons :

- L'étalement d'histogramme ;
- L'égalisation d'histogramme.

II.1.1 L'histogramme d'une image :

Un histogramme est un graphique statistique permettant de représenter la distribution des intensités des pixels d'une image, c'est-à-dire le nombre de pixels pour chaque intensité lumineuse.

D'où l'axe horizontal X indique les différents tons (du sombre au lumineux), l'axe vertical Y indique le nombre de pixel (ou le pourcentage de pixel) pour chaque ton.

Par convention un histogramme représente le niveau d'intensité en abscisse en allant du plus foncé (à gauche) au plus clair (à droite).

Il permet de corriger ou modifier l'exposition de l'image, ainsi que l'échelle des couleurs. Par exemple : améliorer le contraste¹ (rehaussement), corriger une image sous-exposée, renforcer la composante rouge, corriger la non-linéarité du capteur....

Nous trouvons l'histogrammes des images en niveaux de gris qu'il indique pour chaque valeur entre le noir (0) et le blanc (255), et pour les images couleurs, plusieurs histogrammes sont utilisés l'histogramme des luminances et les 3 histogrammes de chacune des composantes R,V,B [11].

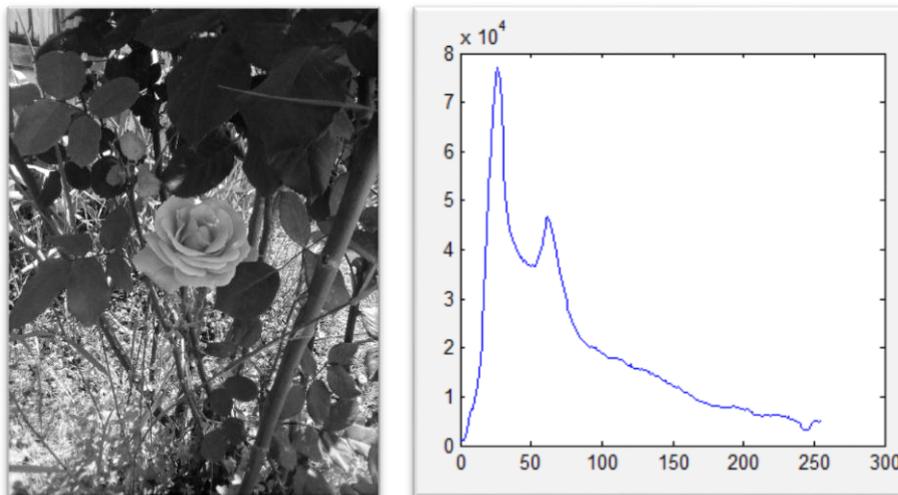


Figure 4: Exemple d'une image et son histogramme.

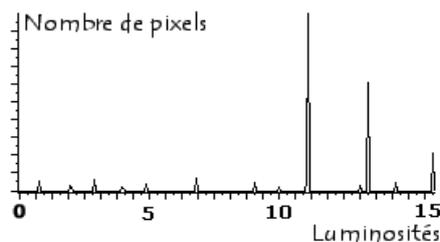


Figure 5: Interprétation d'un histogramme

(<https://www.commentcamarche.com>).

² Propriété intrinsèque d'une image qui quantifie la différence de luminosité entre les parties claires et sombres d'une image.

II.1.2 L'étalement d'histogramme :

L'étalement d'histogramme (aussi appelé "linéarisation d'histogramme" ou "expansion de la dynamique") consiste à répartir les fréquences d'apparition des pixels sur la largeur de l'histogramme. Ainsi il s'agit d'une opération consistant à modifier l'histogramme de telle manière à répartir au mieux les intensités sur l'échelle des valeurs disponibles. Ceci revient à étendre l'histogramme afin que la valeur d'intensité la plus faible soit à zéro et que la plus haute soit à la valeur maximale.

De cette façon, si les valeurs de l'histogramme sont très proches les unes des autres, l'étirement va permettre de fournir une meilleure répartition afin de rendre les pixels clairs encore plus clairs et les pixels foncés proches du noir [12].

Cela consiste à étaler l'image qui est contenue dans l'intervalle [a,b], sur les 256 niveaux de gris disponibles :

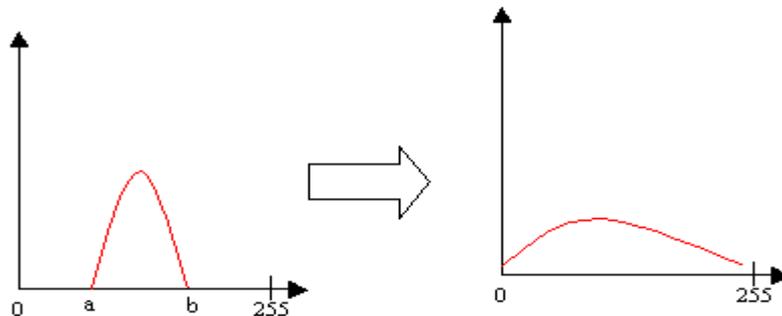


Figure 6: Etalement d'histogramme

(<http://www.tsi.enst.fr>)

On effectue l'étalement d'histogramme avec la transformation linéaire suivante :

$$f'(x, y) = \text{round} \left(M \frac{f(x, y) - N_{\min}}{N_{\max} - N_{\min}} \right) \quad 2.1$$

$f(x, y)$: Niveau de gris au point (x, y) de l'image origine ;

$f'(x, y)$: Niveau de gris au point (x, y) de l'image obtenue ;

N_{\min} : Plus petite valeur de niveau de gris dans l'image origine ;

N_{\max} : Plus grande valeur de niveau de gris dans l'image origine ;

M : représente l'amplitude d'arrivée (en général $M=255$).

Round : est la fonction qui arrondit à l'entier le plus proche.

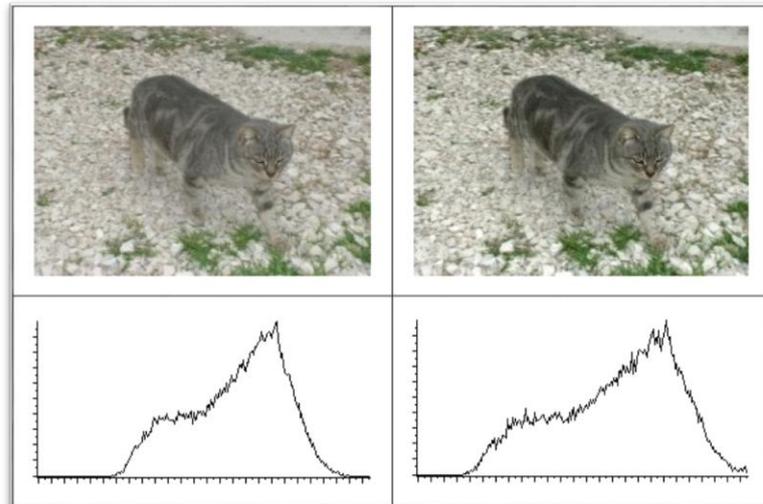


Figure 7: L'étalement d'histogramme

(<https://www.commentcamarche.com>)

A gauche : image en niveaux de gris avec en dessous son histogramme ;
A droite : image obtenue avec l'étalement d'histogramme.

II.1.3 L'égalisation d'histogramme :

L'égalisation d'histogramme est une méthode d'ajustement du contraste d'une image numérique, par lequel on force tous les niveaux de gris de l'image à être équiprobables.

Elle consiste à appliquer une transformation sur chaque pixel de l'image, et donc d'obtenir une nouvelle image à partir d'une opération indépendante sur chacun des pixels. Cette transformation est construite à partir de l'histogramme cumulé de l'image de départ.

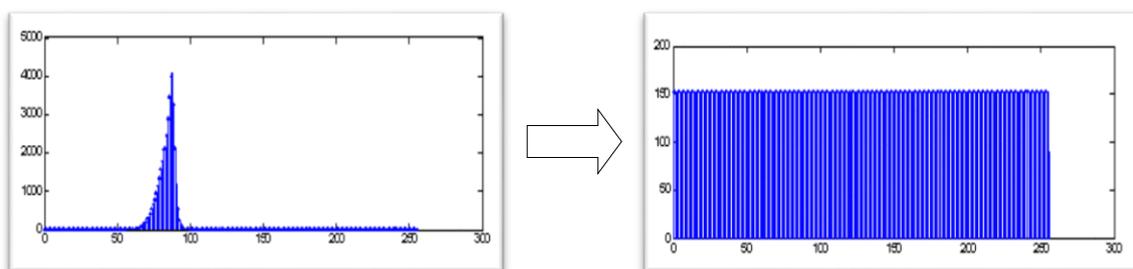


Figure 8: Egalisation d'histogramme

(<http://www.tsi.enst.fr>) .

- Calcul de l'histogramme $h(k)$ avec $k \in [0, 255]$
- Histogramme cumulé $C(k) = \sum_{i=1}^k (h(i))$
- Transformation des niveaux de gris de l'image :

$$I'(x, y) = \frac{C(I(x,y))*M}{N} \quad 2.2 \text{ avec :}$$

- $I'(x, y)$: Niveau de gris au point (x, y) de l'image obtenue ;
- $I(x, y)$: Niveau de gris au point (x, y) de l'image d'origine ;
- N : Nombre total de pixels dans l'image d'origine ;
- M : Amplitude d'arrivée (en général $M=255$).

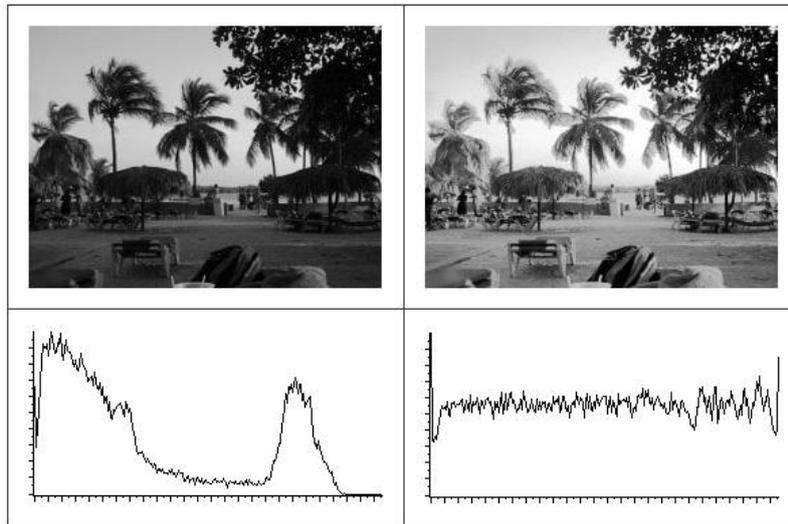


Figure 9: L'égalisation d'histogramme

(<https://www.commentcamarche.com>)

A gauche : image en niveaux de gris avec en dessous son histogramme ;
A droite : image obtenue avec L'égalisation d'histogramme.

III Les filtrages :

Afin d'améliorer la qualité visuelle d'une image, on doit éliminer les effets des bruits (parasites) en lui faisant subir un traitement appelé filtrage d'images.

Le principe du filtrage est de modifier la valeur des pixels d'une image, généralement dans le but d'améliorer la qualité et l'aspect de l'image, par exemple pour l'amélioration du contraste ou un lissage, ou pour détecter les contours.

Le filtrage sert à retrouver le maximum d'informations sous une image bruitée. En pratique, il s'agit de créer une nouvelle image en se servant des valeurs des pixels de l'image d'origine.

On peut catégoriser le filtrage selon trois types [12] :

- Filtrages linéaires ;
- Filtrages non linéaires ;
- Filtrages linéaires itérées ;

III.1 Filtrages linéaires :

Le filtrage linéaire est appelé filtrage linéaire (ou filtrage idéal) car il possède la propriété de linéarité du produit de convolution.

Filtrer une image (x, y) consiste à la convoluer avec une fonction $F(x, y)$ qui est la réponse impulsionnelle du filtre.

Parmi les filtres linéaires usuels dans le traitement d'images sont : le filtre passe-bas et le filtre passe-haut [13].

III.1.1 Le filtre passe-bas :

Le filtre passe-bas sert à diminuer le bruit et atténuer les détails de l'image (augmentation de flou); c'est la raison pour laquelle on parle habituellement de lissage [14].

Parmi les filtres passe-bas : le filtre moyenneur et le filtre gaussien.

- **Le filtre moyenneur :**

Le filtre moyenneur est un filtre linéaire spatial et appelé aussi un filtre de lissage basé sur le principe que la valeur d'un pixel est relativement similaire à son voisinage. Ce filtre consiste à appliquer un masque de convolution h de taille W , puis remplacer la valeur du pixel courant par la moyenne pondérée de ses voisins [15].

L'équation ci-dessous présente un exemple d'un masque de convolution h de taille $W=3$

$$h = \frac{1}{9} \times \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad 2.3$$

La figure suivante présente quelques filtres moyenneurs avec différentes valeurs de voisinage :



a)

b)

c)

d)

Figure 10: Illustration du lissage par filtre moyenneur avec différentes valeurs de voisinage W

(<http://www.tsi.enst.fr>)

a) image originale ; b) image filtrée par $w=3$; c) image filtrée par $w=5$;

d) image filtrée par $w=7$.

- **Le filtre gaussien :**

Le filtre gaussien est aussi un filtre de lissage, il donne un meilleur lissage et une meilleure réduction du bruit que le filtre moyenne, avec le filtre gaussien les contours et les détails fins sont mieux conservés

Une fonction gaussienne est une fonction en exponentielle de l'opposé du carré de l'abscisse (e^{-x^2}). Elle a une forme caractéristique de courbe en cloche.

L'exemple le plus connu est la densité de probabilité de la loi normale (ou loi de Laplace-Gauss ou loi de Gauss [16]).

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad 2.4$$

Notons que :

- la droite $x=m$ est axe de symétrie
- les points d'inflexion sont situés à une distance sigma (σ) de cet axe de symétrie.

Dont la représentation graphique est la suivante :

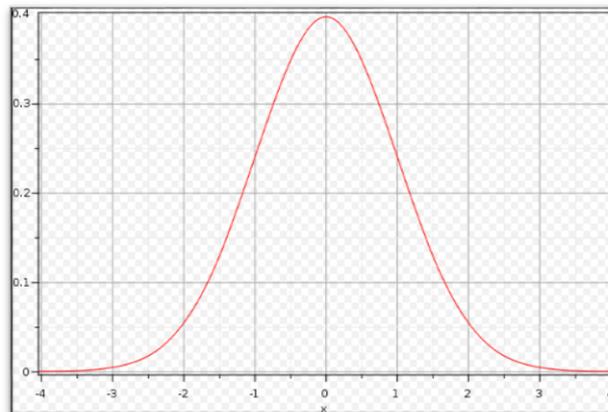


Figure 11: Représentation graphique d'une fonction gaussienne pour $\sigma = 1$

(<https://fr.wikipedia.org>)

Un filtre gaussien est donné aussi par la fonction gaussienne à deux dimensions suivantes :

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad 2.5$$

Où :

σ : est l'écart type;

x, y : valeurs de voisinage.

a) Par exemple si $\sigma = 0,8$; on a le filtre 3x3 suivant :

| | | |
|------------|-----------|-----------|
| $G(-1,-1)$ | $G(0,-1)$ | $G(1,-1)$ |
| $G(-1,0)$ | $G(0,0)$ | $G(1,0)$ |
| $G(-1,1)$ | $G(0,1)$ | $G(1,1)$ |

 $\approx \frac{1}{16}$

| | | |
|---|---|---|
| 1 | 2 | 1 |
| 2 | 4 | 2 |
| 1 | 2 | 1 |

a) Et pour $\sigma = 1$; on a le filtre 5x5 suivant :

 $\approx \frac{1}{300}$

| | | | | |
|---|----|----|----|---|
| 1 | 4 | 6 | 4 | 1 |
| 4 | 18 | 30 | 18 | 4 |
| 6 | 30 | 48 | 30 | 6 |
| 4 | 18 | 30 | 18 | 4 |
| 1 | 4 | 6 | 4 | 1 |

En général un filtre gaussien avec $\sigma < 1$ est utilisé pour réduire le bruit, et si $\sigma > 1$ c'est dans le but de fabriquer une image qu'on va utiliser pour faire un « masque flou ». Il faut noter que plus σ est grand, plus le flou appliqué à l'image sera marqué [17].



Figure 12: Illustration du lissage par filtre gaussien avec différents valeur de voisinage W

La figure suivante montre la différence entre le filtre moyenneur et le filtre gaussien :

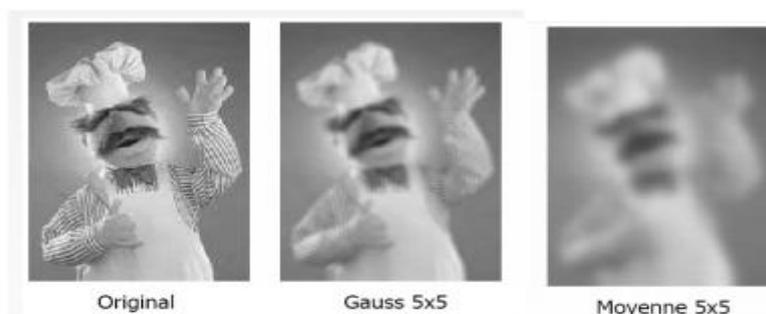


Figure 13: Illustration de lissage par filtre gaussien et filtre moyenneur avec la même valeur de voisinage W =5x5

III.1.2 Le filtre passe-haut :

Le filtre passe-haut permet, d'accentuer les détails et le contour de l'image mais amplifie le bruit ; il s'agit d'un filtre d'accentuation [14].

Parmi les filtres passe-haut on trouve : le filtre sobol et le filtre laplacien

- **Le filtre Sobel :**

Le filtre de Sobel comporte deux masque de taille 3x3 l'un pour l'axe horizontal (X), et l'autre pour l'axe vertical (Y).

Chacun des masques est en fait un filtre gradient, qui sont tous les deux combinés pour créer l'image finale [18].

Les deux filtres de sobol sont :

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

Filtre horizontal

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

Filtre verticale

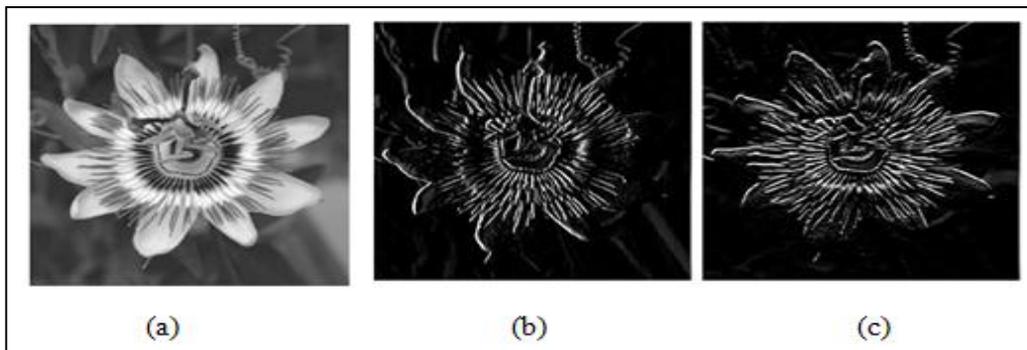


Figure 14: Exemple d'un filtre Sobol

(<https://hal.archives-ouvertes.fr>)

(a) Image originale ; (b) Image horizontale ; (c) Image verticale

- **Le filtre laplacien :**

Le filtre Laplacien est un filtre de convolution particulier utilisé pour mettre en valeur les détails qui ont une variation rapide de luminosité.

Le Laplacien est donc idéal pour rendre visible les contours des objets. Un contour est défini comme une discontinuité locale de l'intensité lumineuse [18].

On peut trouver différents masque Laplaciens de taille 3x3 :

| | | |
|---|----|---|
| 0 | 1 | 0 |
| 1 | -4 | 1 |
| 0 | 1 | 0 |

Laplacien 4-connexe

| | | |
|---|----|---|
| 1 | 1 | 1 |
| 1 | -8 | 1 |
| 1 | 1 | 1 |

Laplacien 8-connexe

| | | |
|----|----|----|
| 1 | -2 | 1 |
| -2 | 4 | -2 |
| 1 | -2 | 1 |

Laplacien de Robinson

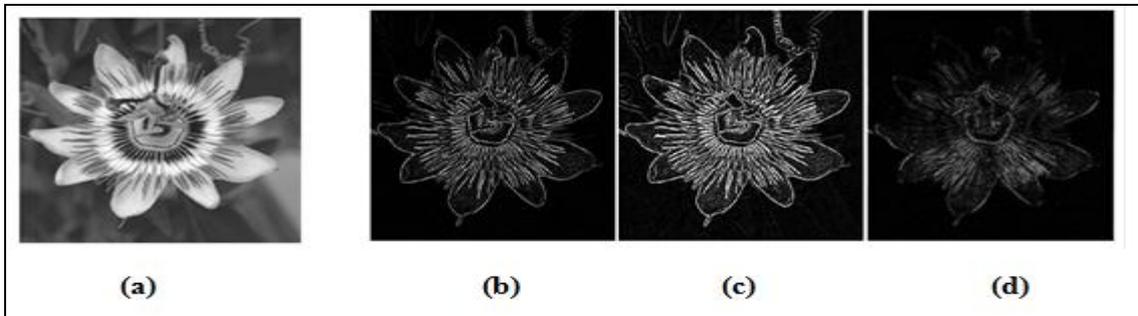


Figure 15: Exemple d'un filtre laplacien

<https://hal.archives-ouvertes.fr>

(a)Image originale ; (b) Laplacien 4-connexe ; (c) Laplacien 8-connexe ;
(d) Laplacien de Robinson

A titre d'exemple on applique à l'image (a) le masque laplacien 8-connexe présentée dans la figure suivante :

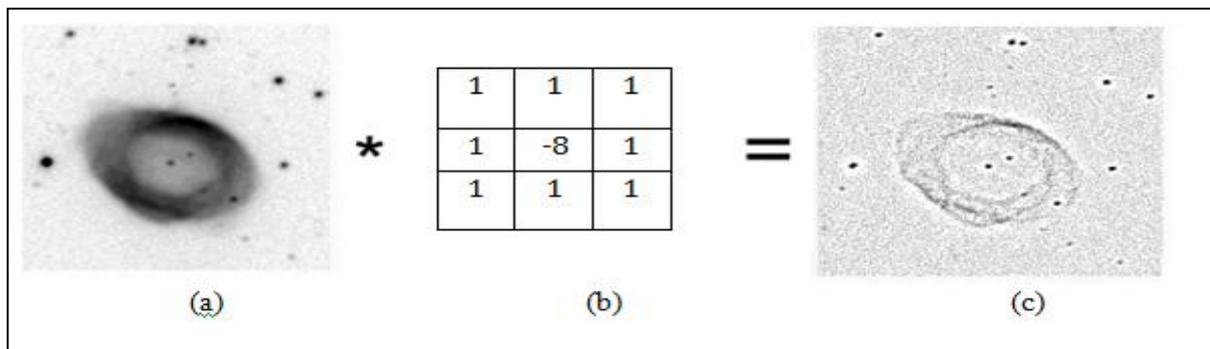


Figure 16: Exemple de masque Laplacien 8-connexe appliqué a une image

(a) Image originale ; (b) masque laplacien 8-connexe ; (c) image résultat ;

III.2 Filtrages non linéaires :

Le filtre non linéaire qui se caractérise par l'élimination du bruit impulsionnel dans l'image et par la conservation des contours des objets dans l'image.

Parmi les filtres non linéaires les plus courants sont : le filtre médian, le filtre MIN et le filtre MAX [15].

III.2.1 Le filtre médian :

Le filtre médian consiste en une opération de lissage non linéaire en parcourant l'image et en remplaçant la valeur du pixel courant par la valeur médiane de ses voisins (définis dans un voisinage de taille W).

Le filtre médian est caractérisé par sa capacité à lisser tout bruit impulsif ou exponentiel, ce qui paraît adapté à notre objectif. Notons que plus la taille du voisinage est grande, plus le lissage est fort et donc plus les détails de l'image sont atténués [15].

D'après la figure suivante, le filtre médian est :

$$F_{med}(W) = F_{med} \{1, 3, 3, 4, 5, 6, 7, 9, 13\} = 5$$

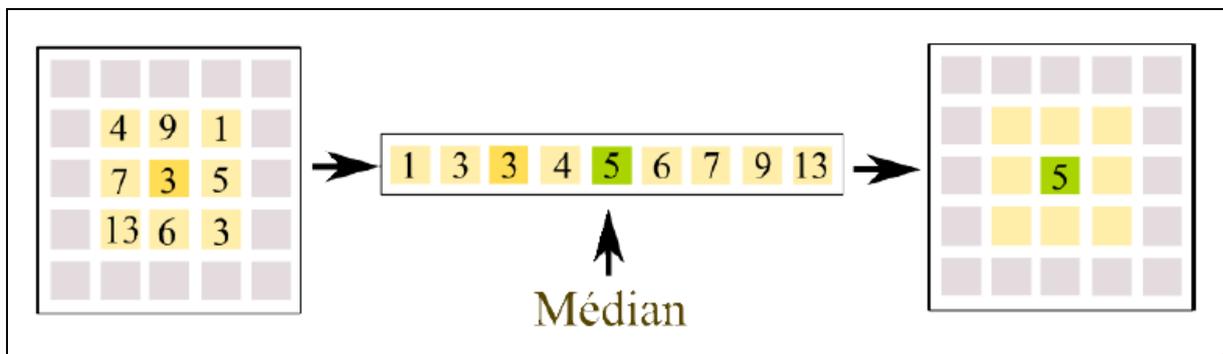


Figure 17 : Illustration du principe du filtre médian avec une matrice de taille $W=3 \times 3$

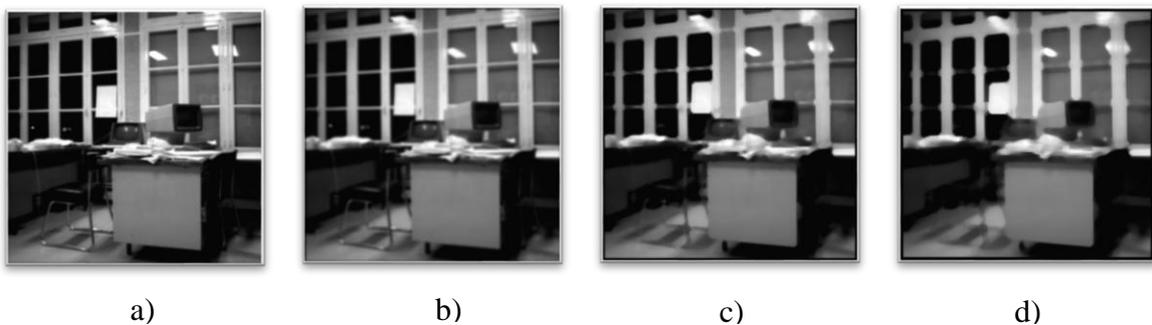


Figure 18: Illustration du lissage par filtre médian avec différentes valeurs de voisinage W (<https://perso.telecom-paristech.fr>)

a) image originale ; b) médian 3×3 ; c) médian 5×5 ; d) médian 7×7 .

III.2.2 Le filtre MIN :

Le filtre MIN remplace la valeur du pixel par la valeur minimale dans son voisinage. Il permet de dilater les pixels sombres de l'image [19].

Si on prend l'exemple précédent, on a :

$W = \{1, 3, 3, 4, 5, 6, 7, 9, 13\}$

Alors le filtre MIN est la valeur minimale de la liste W

Donc $F_{\min}(W) = F_{\min}\{1, 3, 3, 4, 5, 6, 7, 9, 13\} = 1$

III.2.3 Le filtre MAX :

Le filtre MAX remplace la valeur du pixel par la valeur maximale dans son voisinage. Il permet de dilater les pixels clairs de l'image [19].

Si on prend l'exemple précédent, on a : $W = \{1, 3, 3, 4, 5, 6, 7, 9, 13\}$

Alors le filtre MAX est la valeur maximale de la liste W

Donc $F_{\max}(W) = F_{\max}\{1, 3, 3, 4, 5, 6, 7, 9, 13\} = 13$

III.3 Filtrage linéaire itérés :

Il s'agit d'un filtre linéaire amélioré soit :

- En cherchant les meilleures valeurs du masque ;
- En augmentant la taille du masque ;
- En appliquant plusieurs fois de suite le même masque [12].

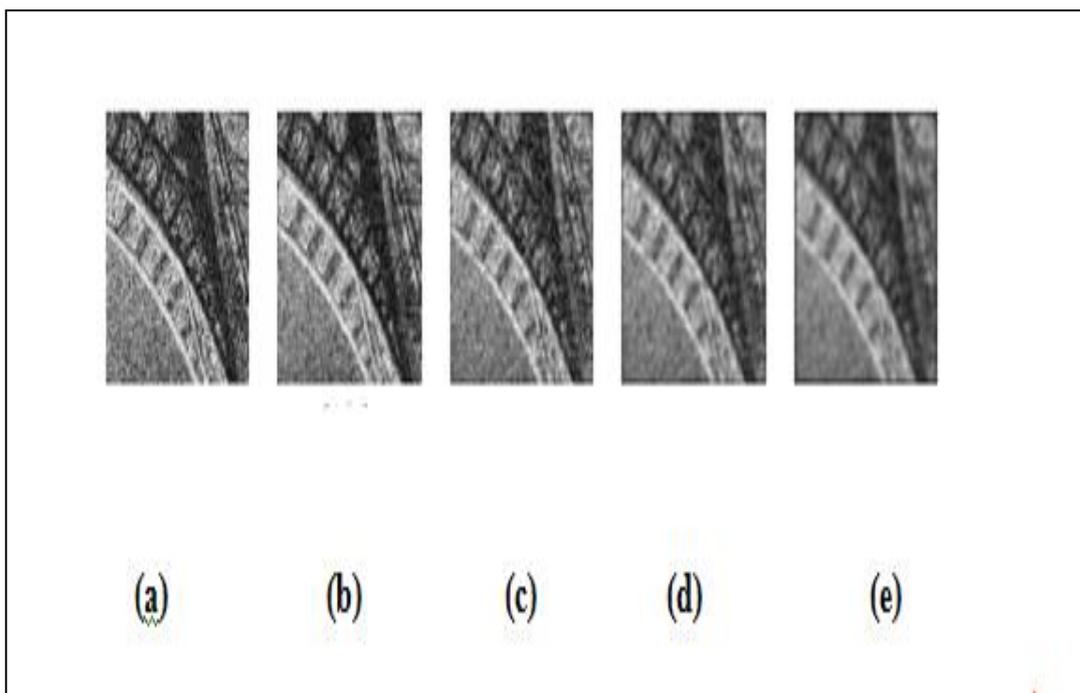


Figure 19: Application du filtre gaussien 3×3 avec $\sigma = 1$:

- (a) image originale ; (b) 1 itération ; (c) 2 itération ; (d) 5 itération
(e) 10 itération.

IV Les opérateurs morpho mathématiques (ou morphologiques) :

La morphologie mathématique est une théorie et technique mathématique et informatique d'analyse de structures. Elle s'applique sur les images qui ont aux niveaux de gris.

L'interprétation d'une image, donne la reconnaissance de certains objets qui se trouvent généralement en deux étapes : la première consiste à repérer sur l'image les structures intéressantes. C'est ce qu'on appelle la *segmentation*, la seconde à *quantifier* ces objets en leur associant des valeurs (nombres ou symboles) en vue de leur classification.

Idée de base de la morphologie mathématique : comparer l'ensemble à analyser avec un ensemble de géométrie connue appelé élément structurant. Un élément structurant est un ensemble qui a les caractéristiques suivantes :

Il possède une forme (géométrie connue), cette forme a une taille repéré par son origine appartenant généralement à l'élément structurant [20].

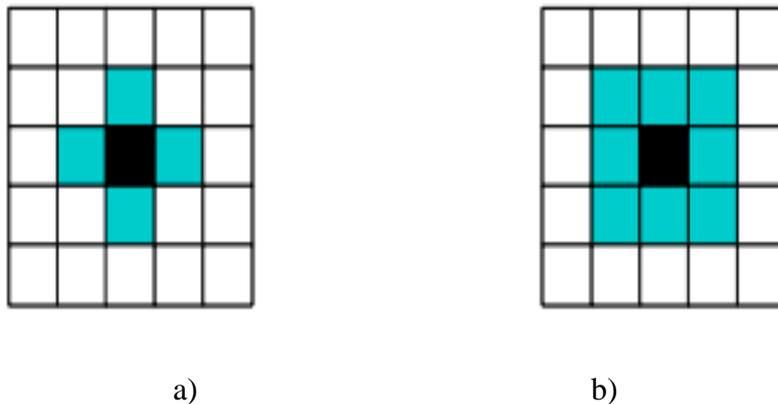


Figure 20: Exemples d'élément structurant

a) Connexité d'ordre 4 ; b) Connexité d'ordre 8

Les deux principales opérations morphologiques sont :

- La dilatation, et
- l'érosion.

On peut trouver d'autres opérations comme :

- L'ouverture : c'est une élimination des éléments fins et modifie les contours ;
- La fermeture : c'est une dilatation qui permet de remplir les petits trous et de lisser les contours ;
- La squelettisation : c'est la représentation la plus fine possible des objets.

L'ouverture et la fermeture ne sont pas inverses car la dilatation et l'érosion ne sont pas inverses.

IV.1 La dilatation :

La dilatation consiste à élargir la figure, la hauteur et largeur de la figure dilatée seront les sommes respectivement des hauteurs et largeurs de la figure originelle et de l'élément structurant [21].

La dilatation d'une forme X par un élément structurant B est notée « $X \oplus B$ ». Elle est définie par :

$$X \oplus B = \{p \in S, \text{tel que } \bar{B}_p \cap X \neq \emptyset\} \quad 2.6$$

Avec « S » le support de l'image, et « p » l'abscisse d'un point P de cette image.

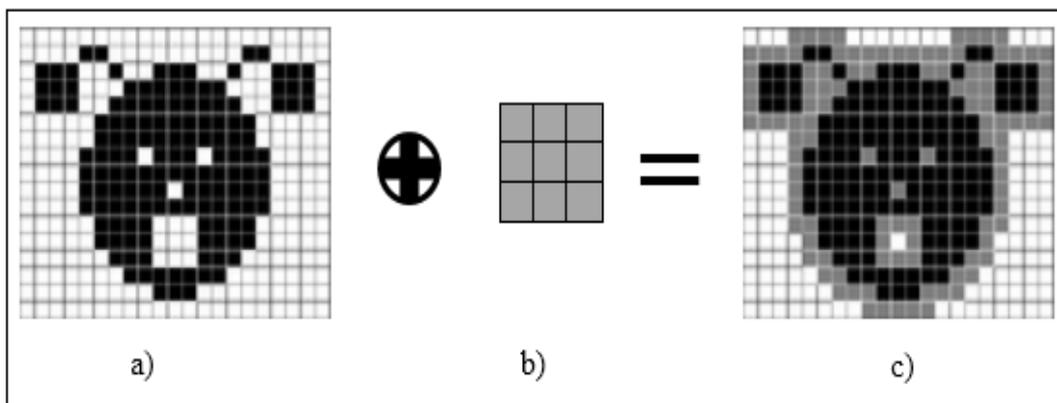


Figure 21: Dilatation d'image

a) Image originale ; b) élément structurant 3*3 ; c) images dilaté ; \oplus : Operateur de dilatation.

IV.2 L'érosion :

L'érosion consiste à rétrécir la figure, la hauteur et largeur de la figure érodée seront les différences respectivement des hauteurs et largeurs de la figure originelle et de l'élément structurant (en particulier si l'élément structurant est plus large ou plus haut que la figure, l'érosion de celle-ci sera vide). Si l'élément structurant est décentré, l'érosion décalera la figure en sens inverse. Enfin les coins concaves de la figure seront déformés en fonction de l'élément structurant [21].

L'érosion d'une forme X par un élément structurant B est notée « $X \ominus B$ ». Elle est définie par :

$$X \ominus B = \{p \in S, \text{tel que } B_p \subseteq X\} \quad 2.7$$

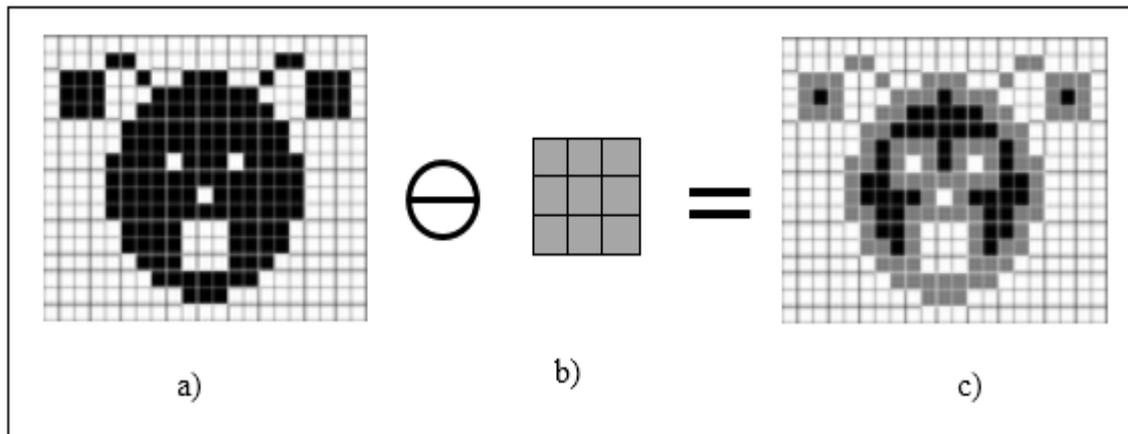


Figure 22: Erosion d'image

a) Image originale ; b) élément structurant 3* 3 ; c) image érodée

\ominus : Operateur d'érosion

V Correction de l'illumination :

Le but de la correction de l'illumination est de supprimer l'hétérogénéité spatiale de l'illumination d'une image causée soit par les défauts du capteur, par une illumination non uniforme de la scène, ou par l'orientation de la surface des objets [22].

La correction de l'illumination est basée sur la soustraction du fond. Ce type de correction suppose que la scène est composée d'un fond homogène et des objets relativement petits et plus clairs ou plus sombres que le fond. Il y a deux techniques de soustraction de fond qui dépendent du fait que le modèle d'illumination de l'image peut être donné par des images supplémentaires qui sont :

- Correction prospective ;
- correction rétrospective ;

V.1 Correction prospective :

La correction prospective utilise des images additionnelles obtenues au moment de l'acquisition de l'image à traiter. Deux types d'images additionnelles peuvent être acquis :

- Une image noire est une image du fond de la scène acquise sans lumière.
- Une image blanche est une image du fond de la scène acquise avec lumière mais sans objet

V.1.1 Correction à l'aide des images noire et blanche :

L'image corrigée $f(x, y)$ (voir fig 21) est obtenue en utilisant la transformation suivante :

$$g(x, y) = \frac{f(x, y) - d(x, y)}{b(x, y) - d(x, y)} * c \quad 2.8$$

Où : $f(x, y)$ est l'image originale, $d(x, y)$ est l'image noire, $b(x, y)$ est l'image blanche, et c est la constante de normalisation qui permet de recouvrer les couleurs de l'image originale :

$$c = \mathit{mean}(f(x, y)) * \frac{1}{\mathit{mean}\left(\frac{f(x, y) - d(x, y)}{b(x, y) - d(x, y)}\right)} \quad 2.9$$

Avec : $\mathit{mean}(f(x, y))$ la valeur moyenne de l'image (x, y) .

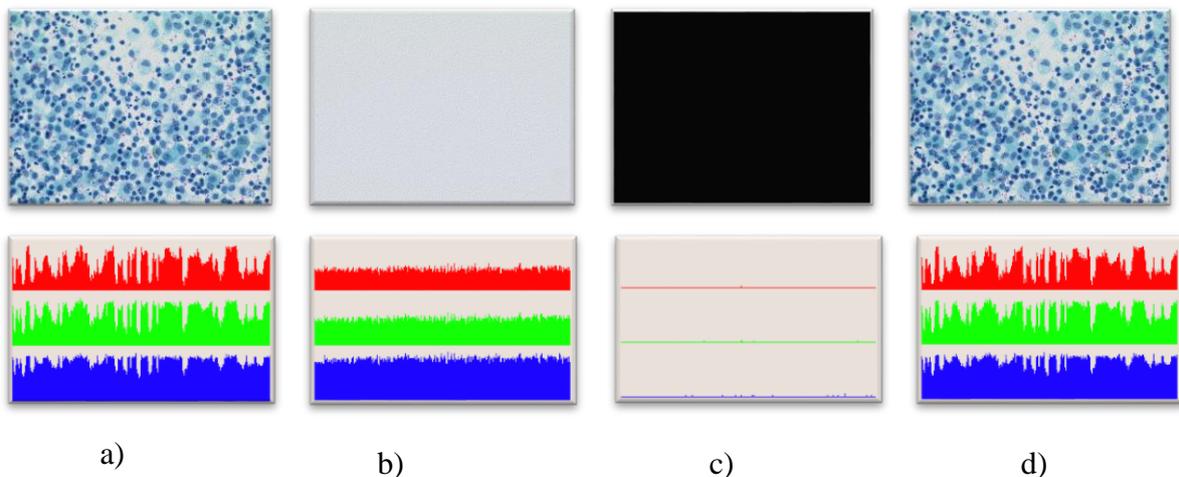


Figure 23: Correction à l'aide des images noire et blanche

- a) Image originale ; b) image blanche ; c) image noire ; d) image résultat ;
au-dessous leurs histogramme associe

Les images sur la seconde ligne représentent les histogrammes, où les niveaux de gris sont présentés sous la forme de barres verticales. On peut constater sur l'image originale que l'éclairage est plus faible à gauche qu'à droite. Ce défaut a été corrigé sur l'image résultat.

V.1.2 Correction à l'aide de l'image blanche seule :

Si seule l'image blanche existe, la méthode utilise une division de l'image source avec l'image blanche si le capteur est linéaire, ou une soustraction de l'image source avec l'image blanche si le capteur est logarithmique avec un gamma de 1.

Dans le cas d'un capteur d'acquisition linéaire, l'image corrigée $g(x, y)$ est obtenue en utilisant la transformation suivante :

$$g(x, y) = \frac{f(x, y)}{b(x, y)} * c \quad 2.10$$

Où $f(x, y)$ est l'image originale, $b(x, y)$ est l'image blanche, et c la constante de normalisation qui permet de recouvrer les couleurs de l'image originale :

$$c = \mathit{mean}(f(x, y)) * \frac{1}{\mathit{mean}(\frac{f(x, y)}{b(x, y)})} \quad 2.11$$

Où : $\mathit{mean}(f(x, y))$ est la valeur moyenne de l'image $f(x, y)$.

Le résultat après les modifications seront l'effet d'éclairage est plus prononcé à gauche qu'à droite.

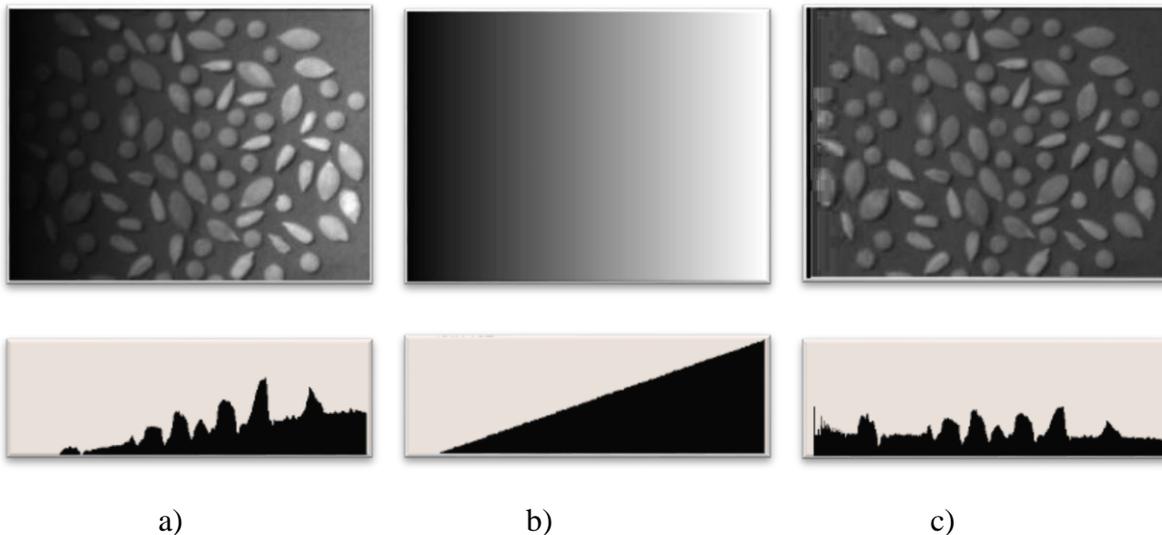


Figure 24 : Correction à l'aide de l'image blanche seule

a) Image originale ; b) image du fond ; c) image résultat ;
au-dessous leurs histogramme associe

V.1.3 Correction à l'aide de l'image noire seule :

Si seule l'image noire existe, la méthode consiste à soustraire l'image noire de l'image originale. Le but est de corriger un défaut lié au capteur. L'image corrigée $g(x, y)$ est alors obtenue en utilisant la transformation suivante :

$$g(x, y) = f(x, y) - d(x, y) + \mathit{mean}(d(x, y)) \quad 2.12$$

Où : $f(x, y)$ est l'image originale $d(x, y)$ est l'image noire et $\mathit{mean}(d(x, y))$ est la valeur moyenne de l'image noire.

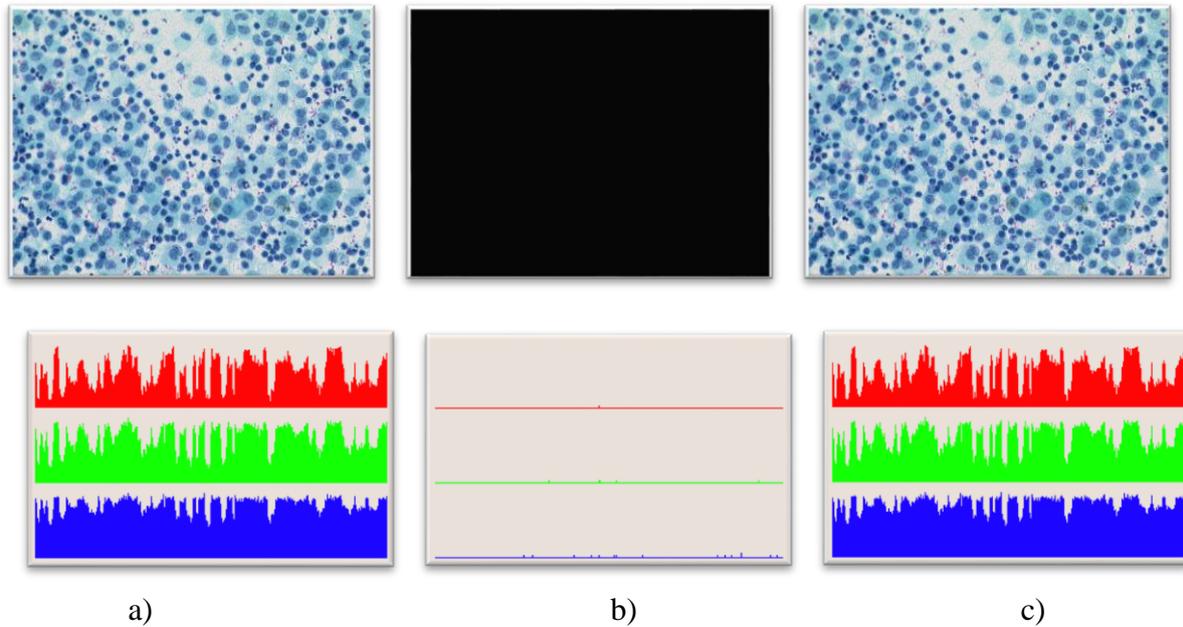


Figure 25 : Correction à l'aide de l'image noire seule

b) Image originale ; b) image noire ; c) image résultat ;
au-dessous leurs histogramme associe

V.2 Correction rétrospective :

Quand les images additionnelles ne sont pas disponibles, un modèle idéal de l'illumination doit être estimé pour construire l'image blanche. La correction rétrospective utilise alors la même technique de suppression du fond que la correction prospective en utilisant cette image blanche estimée.

Il y a différentes techniques pour estimer l'image blanche. Toutes supposent que les objets correspondent aux hautes fréquences et le fond de la scène aux basses fréquences. La correction rétrospective consiste à supprimer les objets du fond pour construire l'image blanche puis à appliquer la même technique que la correction prospective.

V.2.1 Correction rétrospective à base de filtre passe-bas

Le fond d'image est estimé par un filtre passe-bas de grande taille. Le fond estimé est ensuite soustrait de l'image initiale pour compenser le défaut d'illumination. L'image corrigée $g(x,y)$ est obtenue à partir de l'image initiale $f(x,y)$ par :

$$g(x, y) = \left(f(x, y) - LPF(f(x, y)) \right) + mean \left(LPF(f(x, y)) \right) \quad 2.13$$

Où : $LPF(f(x, y))$ est l'image filtrée et $mean(LPf(f(x, y)))$ est la moyenne des valeurs des pixels de l'image filtrée.

L'exemple utilise un filtre gaussien de taille 20, supérieure à la taille des objets (ici les caractères).

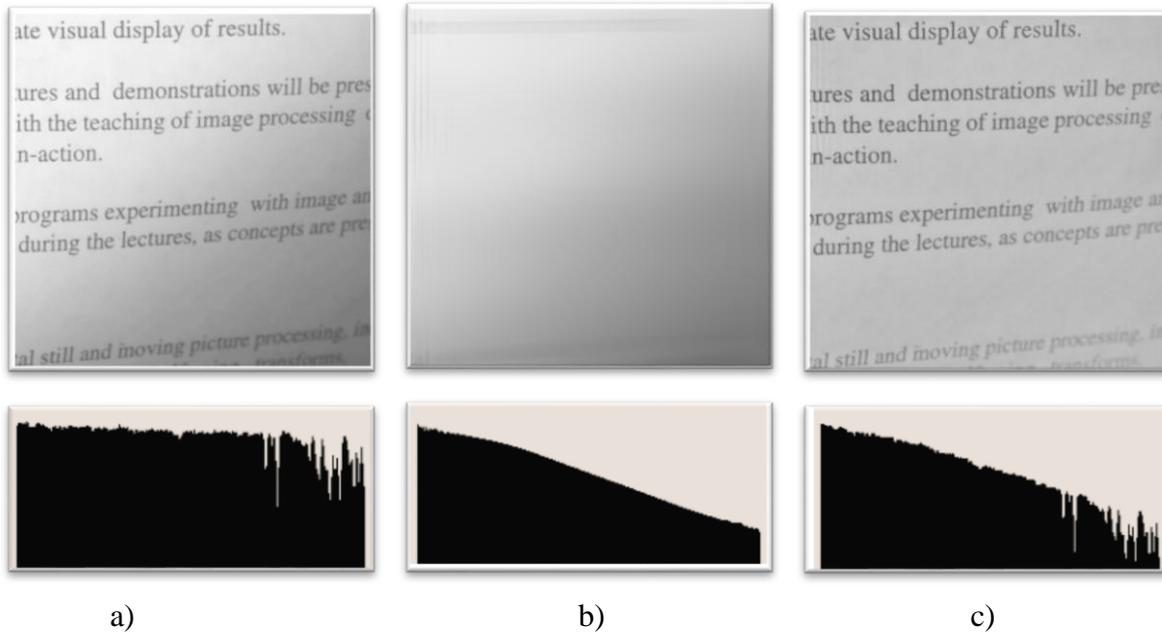


Figure 26 : Correction rétrospective à base de filtre passe-bas

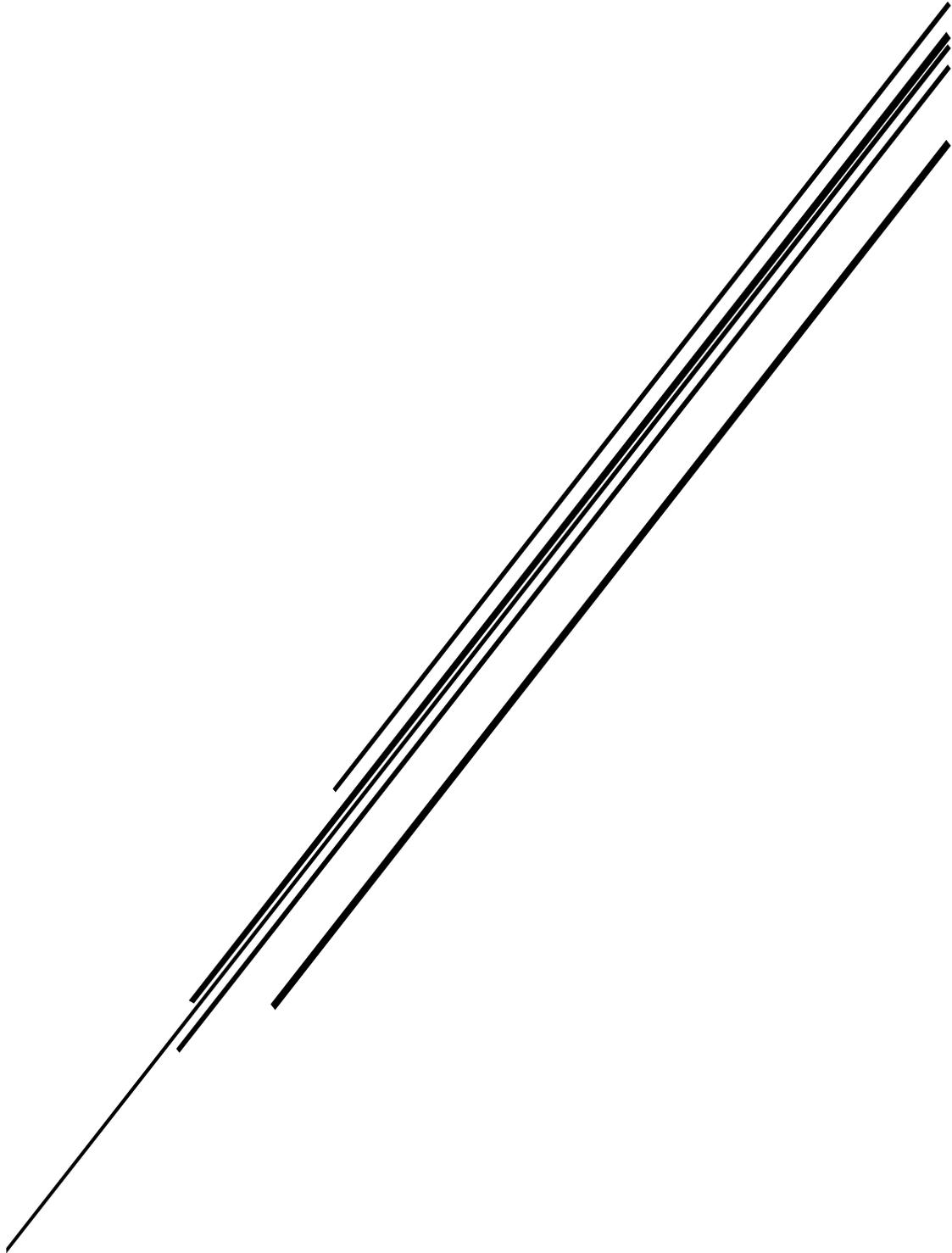
a) Image originale ; b) image du fond estimé ; c) image résultat ;
au-dessous leurs histogramme associe

VI Conclusion :

Dans ce deuxième chapitre, nous avons présenté quelques prétraitements que nous utilisons sur l'image numérique, tel que les modifications de l'histogramme, les filtres ainsi les opérateurs morphologique et à la fin nous avons vu correction de l'illumination .

Nous présentons dans le chapitre suivant la saillance d'image et ses différentes techniques.

Chapitre 3 : SAILLANCE D'OBJET



I Introduction :

De nos jours, le concept de la saillance surgit dans des études consacrées à des phénomènes très variés, tels que la résolution de l'anaphore, la focalisation, la structure informationnelle, etc.

La notion de la saillance peut être appliquée au langage (saillance linguistique), comme elle peut être appliquée au visuel (saillance visuelle). Dans les ouvrages de linguistique, le terme saillance est utilisé pour se référer à l'importance, l'émergence, ou encore la récence d'une entité dans un contexte donné, propriété qui lui permet d'être perçue, repérée plus facilement au milieu d'autres entités. En outre, la saillance visuelle repose sur les mécanismes élémentaires de la vision, qui permettent de modéliser la détection des régions les plus saillantes d'une scène visuelle à partir de ces caractéristiques photométriques et colorimétriques (contrastes de luminance, contrastes colorés, etc.).

II Définition de la saillance :

La notion de saillance contient une idée centrale, mais le détail varie selon le domaine d'application. Un objet est saillant s'il est facilement remarqué. A partir de cette idée intuitive, plusieurs définitions peuvent être proposées, selon le critère qui fait que l'objet va être remarqué.

La saillance, c'est avant tout liée à l'émergence d'une forme sur un fond, la mise en avant d'une entité par rapport à d'autres entités. Etre saillant, c'est ressortir particulièrement, au point de capter l'attention et de donner une accroche [23].

III La détection d'objet saillant :

"Détection d'objet saillant" ou "segmentation d'objet saillant" est généralement interprété dans la vision par ordinateur comme un processus qui comprend deux étapes:

- 1) la détection de l'objet le plus saillant et ;
- 2) segmenter la région précise de cet objet.

Rarement, les modèles distinguent explicitement ces deux étapes. La première étape n'est pas nécessairement besoin d'être limitée à un seul objet. La deuxième étape tombe dans le domaine des problèmes de segmentation classique en informatique vision, mais avec la différence que la précision est ici seulement déterminée par l'objet le plus saillant.

En général, il est convenu que pour une bonne détection de saillance le modèle doit répondre au moins aux trois critères suivants [24] :

- **Bonne détection** : la probabilité de manquer des régions saillantes réelles et faussement marquer l'arrière-plan comme une région saillante devrait être faible.
- **Haute résolution**: les cartes de saillance devraient haute ou pleine résolution pour localiser avec précision les objets saillants et conserver les informations d'image d'origine,
- **Calcul efficacité**: en amont des autres processus complexes, ceux-ci les modèles devraient détecter rapidement les régions saillantes.

IV Les facteurs de la saillance visuelle :

Il existe deux facteurs de la saillance : les facteurs physiques et les facteurs cognitifs [25]:

IV.1 Facteurs physiques :

Il existe plusieurs facteurs physiques en citons :

IV.1.1 La saillance liée aux contrastes de luminosité et de couleur :

Un premier facteur est la saillance intrinsèque à la nature de l'objet en termes de luminosité. Il s'agit par exemple de la saillance du soleil ou d'une ampoule électrique du fait de leurs luminosités inhérentes.

Deuxième facteur que nous proposons, la saillance liée aux variations de perception des couleurs recouvre deux aspects : un premier relatif aux temps de latence, et un second relatif à l'effet spatial des couleurs.

IV.1.2 La saillance liée à une singularité dans un ensemble d'objets :

Il s'agit ici de la saillance due à une mise en relief contextuelle du fait de la présence de plusieurs objets. Compte tenu d'un ensemble d'objets, un élément de cet ensemble est saillant s'il se distingue par une propriété que les autres éléments n'ont pas. Autrement dit, être un singleton est un facteur de saillance.

IV.1.3 La saillance liée à la cohésion et la structuration de la scène :

Un premier facteur dans cette catégorie est la saillance due à la mise en évidence explicite d'un objet dans la scène. Cette mise en évidence peut être due à un éclairage particulier, par exemple dans le cas d'un comédien éclairé par des spots, ou par un contraste figure sur fond important.

Un deuxième facteur très proche mais plus implicite est la saillance due à une construction dédiée. C'est le cas du sujet d'une photographie quand celle-ci est composé de

manière à valoriser ce sujet, par exemple à l'aide d'une construction en triangle qui, plaçant le sujet au sommet supérieur, le rend saillant.

Ensuite, vient la saillance due au placement à un endroit stratégique. Ce facteur fait intervenir la notion de cadre de l'image. Toute image se caractérise par un cadre, dans ce cadre, la centre est un point fort immédiat, c'est-à-dire qu'il attire le regard en priorité. Les autres points forts sont les intersections des lignes horizontales et verticales situées tiers du cadre. Un objet placé à l'un de ces points est potentiellement saillant.

La saillance liée aux répétitions et aux symétries constitue un quatrième facteur. Plusieurs objets identiques, surtout s'ils sont placés de manière régulière, induisent une répétition insistante qui les rend saillants. Dans un même ordre d'idée, plusieurs objets placés de manière symétrique entraînent une perception particulière de ces objets ainsi que celui placé au niveau du centre de symétrie.

Un dernier facteur est la saillance indirecte par transfert visuel de saillance, c'est-à-dire la saillance s'appliquant à un objet qui est reliée à un objet fortement saillant, le lien étant dû à une proximité particulière ou découlant des lignes de force qui dirigent le regard de l'observateur compte tenu de la structure de l'image. Ce facteur s'avère difficile à modéliser, d'une part car l'identification de lignes de force est un problème complexe, et d'autre part car le principe de la transitivité tend à rendre saillants tous les objets de la scène et à enlever tout intérêt à la notion de saillance.

IV.2 Facteurs cognitifs :

La saillance intervient au niveau des processus cognitifs tels que l'intention (volonté du sujet de comprendre le message transmis), l'attention (façon dont un sujet focalise ses capacités de lecture ou d'audition). La mémoire à court terme (en étroite relation avec l'activation, l'accessibilité, la familiarité immédiate, la récence, etc), la mémoire à long terme (l'individu apprend le sens des mots de sa langue à travers des expériences qui lui sont propres et qui connotent des sens particuliers au sein de sa culture et de sa société), la personnalité du destinataire du message et sa situation immédiate (coordonnées spatiales, problèmes, affaire, etc). Chaque sujet met en avant un certain élément du message en fonction de son affect et de ses émotions.

Il existe plusieurs facteurs cognitifs en citons :

IV.2.1 La saillance liée à l'intention :

Un premier facteur cognitif de saillance est ainsi la saillance liée à l'intention, c'est-à-dire à l'influence de la tâche en cours sur la perception. A la suite de son exemple sur les propriétés saillantes d'immeubles, (Edmonds, 1993) [26] souligne que certaines propriétés sont saillantes dans un but précis et pas dans un autre but. Dans son contexte de description d'un itinéraire routier, la propriété de taille est ainsi saillante lorsque la tâche est la désignation d'un immeuble, mais ne l'est lorsqu'il s'agit de désigner une intersection de rues.

IV.2.2 La saillance liée à l'attention :

Un deuxième facteur cognitif, est la saillance liée à l'attention visuelle, dans le sens de focalisation à un sous-espace visuel ou de meilleure détection des stimuli du fait d'une attitude de préparation à la perception. L'attention consiste ces deux cas en un traitement prioritaire de données perceptives.

IV.2.3 La saillance liée à la mémoire à court terme :

Le troisième facteur cognitif, est la saillance liée à la mémoire à court terme, qui concerne les objets récemment traités par le système cognitif visuel,

IV.2.4 La saillance liée à la mémoire à long terme :

Le quatrième facteur cognitif, est la saillance liée à la mémoire à long terme peut se décomposer en deux aspects, premièrement la familiarité visuelle individuelle, et deuxièmement la familiarité visuelle culturelle.

Un exemple de familiarité individuelle est celui de l'infographiste habitué à travailler sur des images et à corriger leurs couleurs : pour lui, un contraste entre deux teintes de rouge sera saillant, alors que le même contraste pourra ne pas être perçu par un autre individu.

Un exemple de familiarité culturelle est le fait que, dans notre vie en société, nous avons l'habitude de communiquer avec nos semblables. Il est ainsi possible de classer selon leur saillance les participants d'une situation visuelle : la présence d'un être humain dans notre champ de vision et saillante avant tout (surtout s'il s'agit de quelqu'un que l'on connaît), vient ensuite la saillance d'un animal, puis celle des objets inanimés.

IV.2.5 La saillance liée à l'affect et aux émotions :

Tout stimulus visuel peut en effet provoquer chez un individu particulier certaines émotions. Par exemple, la perception du visage d'un interlocuteur peut conduire à l'identification immédiate de ce que ressent cet interlocuteur : à partir de l'expression la plus

discrète, du stimulus le plus minime, on est capable de discerner aisément des sentiments tels que la peur, l'amour, la passion, le stress, ou encore le dégoût.

V Les méthodes de la saillance :

Il existe trois méthodes de la saillance, les méthodes basées sur l'estimation du contraste local, les méthodes basées sur l'estimation du contraste global et les méthodes hybrides qui utilisent une estimation de la saillance basée sur le contraste global et le contraste local à la fois [27].

V.1 Méthodes basées sur l'estimation du contraste local :

Ces méthodes sont généralement liées à la prédiction des fixations du regard en explorant la rareté des régions d'images par rapport à leur voisinage. Par exemple, (Hu et al) [28] utilisent la différence de texture entre une région et ses régions voisines comme critère pour estimer le contraste local. (Gao et al) [29] ont proposé une méthode mesurant la saillance comme le pouvoir discriminatif des caractéristiques utilisées d'une région par rapport aux régions voisines. Ils utilisent l'information mutuelle pour estimer le contraste local.

V.2 Méthodes basées sur l'estimation du contraste global :

Ces méthodes mesurent la rareté et l'unicité d'une région par rapport à l'ensemble de l'image. Plusieurs types d'approches ont été proposés pour mesurer le contraste global d'une région de l'image à traiter. La manière la plus triviale est d'utiliser la théorie des probabilités et statistiques en assignant une valeur de saillance élevée aux régions irrégulières.

L'information à priori sur la compacité de couleur a été largement utilisé dans ce contexte. Cela suppose que plus une couleur possède une distribution de couleur élevée dans l'image, moins elle est susceptible d'être contenue dans l'objet saillant. Dans ce contexte, l'utilisation de la théorie des probabilités semble être un outil efficace. Outre les probabilités, la distance colorimétrique entre régions, la théorie des graphes, la distance colorimétrique et l'utilisation du domaine fréquentiel ont été largement utilisés dans ce contexte.

V.2.1 Méthodes probabilistes :

Les méthodes basées sur les probabilités sont efficaces dans les cas de la présence d'un seul objet saillant dominant dans l'image mais peuvent faillir sensiblement dans le cas de scènes complexes spécialement lorsque l'arrière-plan possède une distribution colorimétrique similaire à celle de l'objet saillant.

V.2.2 Méthodes graphiques : on peut citer deux méthodes graphiques :

V.2.2.1 Méthode graphique basée sur les marches aléatoires markoviennes :

Comme réponse aux limites des méthodes probabilistes, (Harel et al) [30] proposent un modèle basé sur une représentation graphique pour la détection de la saillance. L'image est représentée sous forme d'un graphe non orienté et pleinement connecté dont les nœuds correspondent à des blocks et les arcs sont des liens pondérés en fonction du degré de similarité et la distance spatiale entre les nœuds. Les marches aléatoires sont appliquées sur le graphe dont la transition entre un nœud et un autre est contrôlé par les liens du graphe. Durant le processus de marches aléatoires, les nœuds les moins visités sont considérés les plus saillants dû à leur unicité ou rareté dans un contexte global. Nous supposons que $\{\mathcal{F}_k\}_{k=1}^K$ représente l'ensemble des caractéristiques utilisées. Pour la $k^{ième}$ caractéristique, le degré de similarité entre deux nœuds positionnés dans les emplacements (i, j) et (p, q) respectivement est défini comme suit:

$$d_k((i, j) \parallel (p, q)) = \left| \log \frac{\mathcal{F}_k(i, j)}{\mathcal{F}_k(p, q)} \right| \quad 3.1$$

Notons que dans certains cas, la similarité logarithmique peut être remplacée par une simple différence $\mathcal{F}_k(i, j) - \mathcal{F}_k(p, q)$ qui est une opération plus simple avec des performances comparables.

V.2.2.2 Méthode graphique basée sur la distance géodésique :

Dans [31], un modèle graphique basé sur la distance géodésique est proposée pour estimer la saillance. Les auteurs considèrent un template noté B défini à partir des contours de l'image comme information à priori sur l'arrière-plan. Ils effectuent ensuite une segmentation de l'image en régions sur laquelle ils construisent un graphe $G = (\mathcal{V}, \mathcal{E})$ pleinement connecté unidirectionnel où l'ensemble des nœuds (\mathcal{V}) correspond aux régions de l'image et au template B comme région virtuelle ($\mathcal{V} = \{r_i\} \cup B$) et \mathcal{E} dénote l'ensemble des arcs du graphe.

Il existe deux types d'arcs dans le graphe: les arcs internes qui relient les régions adjacentes et les arcs du contour de l'image qui relient les nœuds du contour à B .

$$\mathcal{E} = \{(r_i, r_j) | r_i \text{ est adjacent à } r_j\} \cup \{(r_i, B) | r_i \text{ est sur le contour de l'image}\} \quad 3.2$$

Un arc du graphe est pondéré en fonction de la distance géodésique entre les deux nœuds qu'il relie (r_i et r_k). Il s'agit de la somme des poids sur le plus court chemin reliant r_i à r_k défini comme suit:

$$d_{geo}(r_i, r_k) = \min_{r_{i1}=r_i \dots r_{in}=r_k} \sum_{j=1}^{n-1} D_r(r_{ij}, r_{i,j+1}) ; (r_{ij}, r_{i,j+1}) \in \mathcal{E} \quad 3.3$$

Le score de saillance de r_i est considéré comme sa distance géodésique par rapport à B sur le graphe G comme suit:

$$s(r_i) = d_{geo}(r_i, B) \quad 3.4$$

V.2.3 Méthodes basées sur la distance colorimétrique :

D'autres travaux ont eu recours à la distance colorimétrique pour estimer le contraste d'une région par rapport à toutes les autres régions de l'image. Dans ce contexte, (M.M. Cheng et al) [32], proposent un modèle basé sur la distance spatiale et la distance entre histogrammes de couleur des différentes régions de l'image pour estimer la saillance.

D'abord, l'image est segmentée en régions. La saillance pour une région donnée r_k est calculée par rapport à toutes les autres régions de l'image comme suit :

$$S(r_k) = \sum_{r_k \neq r_i} w(r_i) D_r(r_k, r_i) \quad 3.5$$

Avec : $D_r(r_k, r_i)$ est la distance colorimétrique et $w(r_i)$ correspond au nombre de pixels dans r_i qui favorise les grandes régions à avoir une saillance importante.

La distance colorimétrique entre deux régions r_1 et r_2 est calculée comme suit :

$$D_r(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f(c_{1,i}) f(c_{2,j}) D(c_{1,i}, c_{2,j}) \quad 3.6$$

Avec : $f(c_{k,i})$ la probabilité de la $i^{\text{ème}}$ couleur $c_{k,i}$ parmi toutes les n_k couleurs de la $k^{\text{ème}}$ région r_k , $k = \{1,2\}$.

V.2.4 Méthodes d'estimation de la saillance dans le domaine fréquentiel :

Certaines autres méthodes proposent d'estimer la saillance globale dans le domaine fréquentiel, (Hou et Zhang) [33] considèrent que les similarités entre régions impliquent des redondances dans l'image et développent un modèle de saillance basé sur le calcul du résidu spectral. Ils supposent que les singularités statistiques dans le spectre peuvent être responsables de régions anormales dans l'image, où les objets deviennent visibles.

D'abord, ils dérivent l'amplitude $\mathcal{A}(f)$ et la phase $\mathcal{P}(f)$ à partir d'une image $I(x)$ en appliquant la transformée de Fourier (\mathcal{F}). Le log-spectrum $\mathcal{L}(f)$ est ensuite calculé à partir de l'amplitude.

$$\mathcal{A}(f) = \mathcal{R}(\mathcal{F}[I(x)]), \quad \mathcal{P}(f) = \varphi(\mathcal{F}[I(x)]), \quad \mathcal{L}(f) = \log(\mathcal{A}(f)) \quad 3.7$$

A partir de $\mathcal{L}(f)$, le résidu spectral ($\mathcal{R}(f)$) peut être obtenu en multipliant $\mathcal{L}(f)$ avec un filtre $n \times n$ local ($h_n(f)$) puis en soustrayant le résultat de $\mathcal{L}(f)$ comme suit:

$$\mathcal{R}(f) = \mathcal{L}(f) - h_n(f) * \mathcal{L}(f) \quad 3.8$$

En utilisant la transformée de Fourier inverse, ils construisent la carte de saillance S dans le domaine spatial comme suit:

$$S(x) = g(x) * \mathcal{F}^{-1}[\exp(\mathcal{R}(f) + \mathcal{P}(f))]^2 \quad 3.9$$

Avec : $g(x)$ un filtre gaussien qui assure un meilleur effet visuel.

(Achanta et al) [34], ont implémenté une approche dans le domaine fréquentiel pour la détection de régions saillantes en utilisant les caractéristiques de couleur et de luminance dans l'espace de couleur CIE Lab. La saillance en un point x est estimée comme la différence (distance euclidienne) entre le vecteur représentant la moyenne des caractéristiques de l'image I_μ et le vecteur des caractéristiques de x après lissage de l'image par un filtre gaussien comme suit:

$$s(x) = \|I_\mu - I_w(x)\|^2 \quad 3.10$$

Où : I_w représente le résultat du lissage de l'image I .

V.3 Méthodes hybrides :

Comme solution aux inconvénients des méthodes basées sur l'estimation du contraste local et les méthodes basées sur l'estimation du contraste global, plusieurs méthodes proposent de combiner les informations issues des deux types de contraste pour la détection d'objets saillants. Parmi ces méthodes, nous pouvons distinguer d'autres types d'approches: méthodes probabilistes, méthodes basées sur les graphes et les autres méthodes :

V.3.1 Méthodes probabilistes :

La saillance d'un objet pouvant être définie comme sa singularité par rapport aux autres, il est assez naturel d'utiliser une théorie probabiliste ou statistique afin de relier la saillance aux éléments les moins probables d'une scène ou d'une partie de la scène.

V.3.2 Méthodes basées sur les graphes : il existe deux méthodes :

V.3.2.1 Méthodes basées sur une seule résolution :

Ces méthodes se basent généralement sur un graphe mono-couche construit sur la base d'une segmentation de l'image en régions (généralement des superpixels) qui sont considérés comme les nœuds du graphe :

V.3.2.1.1 Marches aléatoires markoviennes :

Dans [35], les auteurs proposent un modèle basé sur les marches aléatoires markoviennes pour détecter les objets saillants construit sur un graphe mono-couche. Ils segmentent d'abord l'image en superpixels, puis construisent un graphe $G(V, E)$ où les nœuds V sont représentés par les superpixels et les arcs E des liens pondérés reliant les nœuds. Comme les objets saillants occupent rarement les bords de l'image, les auteurs considèrent les nœuds qui touchent les bords comme absorbants (arrière-plan virtuel). Notons que les arcs relient uniquement chaque nœud à ses voisins transitoires. Les nœuds absorbants ne sont donc pas interconnectés.

V.3.2.1.2 Connectivité avec l'arrière-plan :

Dans [36], les auteurs segmentent l'image à traiter en superpixels, puis définissent un graphe non orienté dont chaque nœud correspond à un superpixel. Le poids reliant deux nœuds adjacents du graphe est défini comme la distance euclidienne d_{euc} entre leur moyenne colorimétrique dans l'espace CIE LAB. La distance géodésique entre deux nœuds p et q est calculée comme la somme des poids du chemin le plus court qui les sépare comme suit:

$$d_{geo}(p, q) = \min_{p_1=p, p_2, \dots, p_n=q} \sum_{i=1}^{n-1} d_{euc}(p_i, p_{i+1}) \quad 3.11$$

Nous notons que cette distance a été utilisée dans [37] comme seule mesure pour estimer la saillance dans un graphe. A partir de cette mesure, les auteurs définissent la zone que couvre l'influence d'un nœud p comme suit:

$$Area(p) = \sum_{i=1}^N \exp\left(-\frac{d_{geo}^2(p, p_i)}{2\sigma^2}\right) = \sum_{i=1}^N S(p, p_i) \quad 3.12$$

Avec : N le nombre de nœuds du graphe et σ un paramètre de stabilisation fixé empiriquement à 10. Ils définissent ensuite une mesure basée sur les informations à priori sur la connectivité et l'arrière-plan pour estimer le taux d'overlapping (connectivité) d'une région avec les contours de l'image. Celle-ci est calculée comme suit:

$$Over = \frac{\sum_{i=1}^N S(p, p_i) \cdot \delta_i}{\sqrt{Area(p)}} \quad 3.13$$

Avec : $\delta_i = 1$ si le nœud i se trouve sur un des bords de l'image, et $\delta_i = 0$ sinon. A partir de cette mesure, ils formalisent la connectivité à l'arrière-plan pour chaque superpixel i sous forme d'une probabilité comme suit:

$$w_i^{bg} = 1 - \exp\left(-\frac{Over^2(p_i)}{2\varepsilon^2}\right) \quad 3.14$$

Avec : ε une constante fixée empiriquement à 1. Le contraste d'un nœud est alors calculé en fonction de sa distance euclidienne (d_{euc}), sa distance géodésique par rapport à tous les autres nœuds et sa connectivité par rapport à l'arrière-plan comme suit:

$$w_j^{fg} = \sum_{i=1}^N d_{euc}(p_j, p_i) \cdot d_{geo}(p_j, p_i) \cdot w_i^{bg} \quad 3.15$$

Un post-processus d'optimisation est alors effectué afin d'améliorer le contraste entre les objets saillants et l'arrière-plan en minimisant la fonction suivante:

$$\sum_{i=1}^N w_i^{bg} s_i^2 + \sum_{i=1}^N w_i^{fg} (s_i - 1)^2 + \sum_{i,j} w_{ij} (s_i - s_j)^2 \quad 3.16$$

Avec : $\{s_i\}_{i=1}^N$ les valeurs de saillance des N nœuds du graphe.

V.3.2.2 Méthodes basées sur plusieurs résolutions :

Ces méthodes se basent généralement sur un graphe multicouche ayant une structure hiérarchique, on peut citer :

V.3.2.2.1 Structure arborescente pour la détection de la saillance :

Dans [38], les auteurs segmentent d'abord l'image à traiter en n régions primitives en se basant sur une méthode de détection de contours [39]. Ils calculent ensuite la similarité régionale entre chaque paire de segments R_i et R_j .

A partir de la mesure de similarité régionale, les auteurs estiment ensuite le niveau de saillance de chaque région en se basant sur trois mesures: le contraste global, la dispersion spatiale et la connectivité. Le contraste global est défini pour une région R_i comme la différence de couleur pondérée avec toutes les autres régions de l'image selon l'équation suivante:

$$GC(R_i) = \sum_{j=1}^n |R_j| \cdot Sim_d(R_i, R_j) \cdot \|mc_i - mc_j\| \quad 3.17$$

Avec: mc_i dénote la moyenne de couleur de la région R_i . Le poids $|R_j| \cdot Sim_d(R_i, R_j)$ indique que les régions larges et proches de R_i possèdent une meilleure contribution dans l'évaluation du contraste globale de R_i . La normalisation du contraste globale de R_i se fait comme suit:

$$NGC(R_i) = \frac{GC(R_i) - GC_{min}}{GC_{max} - GC_{min}} \quad 3.18$$

Avec : GC_{min} et GC_{max} sont les valeurs minimums et maximums respectivement dans la mesure du contraste de toutes les régions de l'image. Comme les régions ayant des valeurs de similarités régionales similaires doivent avoir des valeurs de contraste global similaires.

V.3.2.2.2 Structure hiérarchique pour la détection de la saillance :

Dans [40], un modèle de détection d'objets saillants basé sur un graphe hiérarchique à plusieurs résolutions a été proposé. Les auteurs définissent une structure de trois résolutions superposées organisés de bas en haut selon l'ordre décroissant des granularités des différentes segmentations. La couche la plus basse est construite à partir d'un processus de fusion des régions adjacentes et similaires d'une carte de segmentation à forte granularité (Over-segmentation). Chacune des autres résolutions est obtenue en appliquant le même processus de fusion sur la résolution se trouvant au niveau inférieur. Une carte de saillance intermédiaire est calculée ensuite pour chaque résolution.

La figure suivante présente les différentes résolutions :



Figure 27: Exemple illustrant les différentes résolutions du graphe proposé dans [17].

(a) Image en couleur ; (b) Over-segmentation ; (c) Résolution 1 ; (d) Résolution 2 ;
(e) Résolution 3

V.3.2.2.3 Multi-graphes pour la détection d'objets saillants :

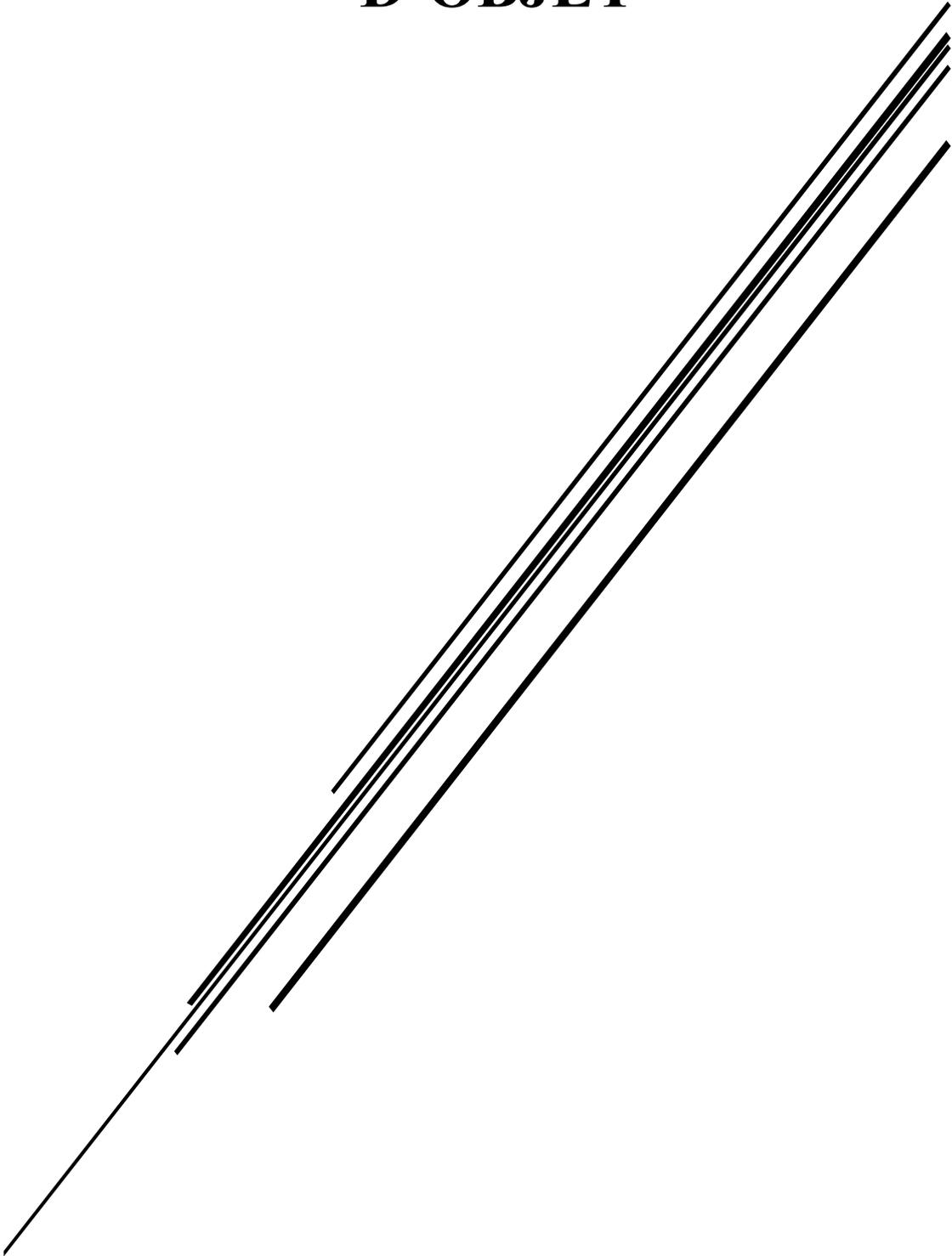
Dans [39], les auteurs ont utilisé les principes de la physique quantiques sur une structure multi-graphes pour estimer la saillance des objets. Ils définissent trois résolutions pour l'image à traiter. Les différentes résolutions sont obtenues en segmentant l'image à traiter en superpixels avec des niveaux de granularité différents. Pour chaque résolution, les auteurs construisent un graphe où les superpixels représentent les nœuds. Ils définissent un arc pondéré non orienté entre chaque paire de nœuds adjacents. Les auteurs considèrent que deux nœuds sont m -voisins s'il est possible d'atteindre un nœud à partir de l'autre en parcourant $m-1$ arcs

VI Conclusion :

Dans ce chapitre, nous avons fait un aperçu général sur la notion de saillance qui varie selon le domaine d'application, afin d'exploiter de façon optimale, les informations qu'elle contient. Nous avons aussi vu les différentes méthodes et facteurs de saillance.

Dans ce qui suit nous allons focaliser sur la segmentation d'objet, et étudierons leurs différentes techniques et méthodes.

Chapitre 4 : SEGMENTATION D'OBJET



I Introduction :

La segmentation est une étape primordiale en traitement d'images. Elle constitue un problème fondamental en vision par ordinateur et en traitement d'image. Elle est utilisée dans de nombreuses applications étroitement liées à la vie quotidienne de l'être humain comme dans le domaine de la biomédecine, la détection des objets et le suivi de leurs mouvements, la reconnaissance faciale ou de l'empreinte digitale, etc. Elle peut être définie comme une tâche de séparation des objets du fond dans les images.

L'objet principal d'un processus de segmentation d'objet est l'extraction d'un ensemble d'attribut significatif permettant de caractériser les entités présentes dans les images.

II Définition de la segmentation :

En traitement d'image, la segmentation est un procédé qui a pour but de rassembler les pixels de l'image étudiée en région selon des critères prédéfinis [41].

La segmentation consiste à partitionner une image en plusieurs sous-ensembles appelés régions. Une région est un ensemble connexe de pixels ayant des propriétés communes (intensité, texture...) qui les différencient des pixels des régions voisines.

De nombreux critères de segmentation existent ; suivant le domaine d'application et le type d'images traitées, le critère prendra en compte le niveau de gris, la texture, la couleur, etc [42].

III Méthodes de segmentation :

Il existe différentes méthodes de segmentation, nous citons entre autres :

- La méthode de segmentation par seuillage (la binarisation) ;
- La méthode de segmentation par croissance de régions ;
- La méthode de segmentation par les contours ;
- Les méthodes de décomposition/ fusion (split and merge) ;
- Graph cut.

III.1 La méthode de segmentation par seuillage (la binarisation) :

Segmenter une image en couleurs peut être extrêmement coûteux. Afin de simplifier les traitements, de nombreuses recherches se sont orientées vers la binarisation de l'image.

Binariser une image appelée aussi seuillage, est la technique de classification la plus simple. Les pixels de l'image sont partagés (ou segmentés) par un seuil T en deux classes : le fond et l'objet. Ensuite, binariser une image permet aussi de compresser les documents de manière plus importante [43].

La binarisation permet alors de conserver l'information comprise entre 0 et T ou entre T+1 et 255.avec cet algorithme :

$$B(x, y) = \begin{cases} 0 & \text{si } I(x, y) < T \text{ (noir)} \\ 1 & \text{sinon (blanc)} \end{cases} \quad 4.1$$

Généralement la binarisation s'opère sur des images en niveaux de gris. Pour passer d'une image couleurs à une image en niveaux de gris on peut appliquer ces formules :

(a) Formule standard :

$$\text{Gris} = \frac{\text{Rouge} + \text{Vert} + \text{Bleu}}{3} \quad 4.2$$

(b) Formule proposée par la C.I.E. (Commission Internationale de l'éclairage) dans sa recommandation 709, qui concerne les couleurs « vraies » ou naturelles :

$$\text{Gris} = 0.2125 * \text{rouge} + 0.7154 * \text{Vert} + 0.0721 * \text{Bleu} \quad 4.3$$

Il existe deux types de seuillage :

III.1.1 Seuillage local :

Dans cette approche, le seuil n'est plus unique mais est déterminé pour chaque pixel ou pour des régions de pixels aux caractéristiques voisines. Le seuil est calculé automatiquement pour chaque pixel de l'image. Les méthodes de Sauvola ou de Niblack sont utilisées pour la binarisation locale.

III.1.2 Seuillage global :

Le principe du seuillage global est d'utiliser une valeur seuil à partir de laquelle on peut choisir à quelle classe le pixel appartient.

Le seuillage globale peut être manuelle ou automatique. La méthode manuelle présente un inconvénient car le seuil optimal dépend d'une image à une autre. La méthode automatique est meilleure mais elle peut échouer si l'image est dégradée (mauvais éclairage, bruits...).

Le tableau 1 indique les principes et les inconvénients de seuillage local et global

| Type | Principe | Inconvénients |
|--------------------------------------|--|--|
| Seuillage global (Otsu, 1979) | D'après l'histogramme, cherche à maximiser la variance intra-classe. | Problèmes pour les documents mal éclairés. |

| | | |
|--|---|---|
| Seuillage local (Bernsen, 1986) | Estime la valeur du seuil en faisant la moyenne de la plus haute et la plus basse valeur de la fenêtre. | Le seuil est trop bas lorsque la fenêtre est centrée sur du fond. |
| Seuillage local (Niblack, 1986) | Amélioration de Bernsen : prise en compte de la variance et de la moyenne. | Même problème que Bernsen: apparition de bruit sur les zones uniformes. |
| Seuillage local (Trier, 1995) | « Ternarise » l'image en fonction du gradient puis utilise une heuristique pour réduire à deux classes. | L'utilisation de seuils définis empiriquement rend la méthode peu robuste. |
| Seuillage global (Wu, 1998) | « Floute » l'image pour mieux séparer l'histogramme et utilise une méthode de seuillage global Otsu. | Problèmes lorsqu'il n'y a pas deux modes distincts sur l'histogramme. |
| Seuillage local (Sauvola, 2000) | Insère des constantes dans la méthode de Niblack afin d'améliorer la méthode sur les zones uniformes | Les constantes à ajuster empêchent la méthode de traiter parfaitement des documents non uniformes |

Table 1: Différentes techniques de binarisation.

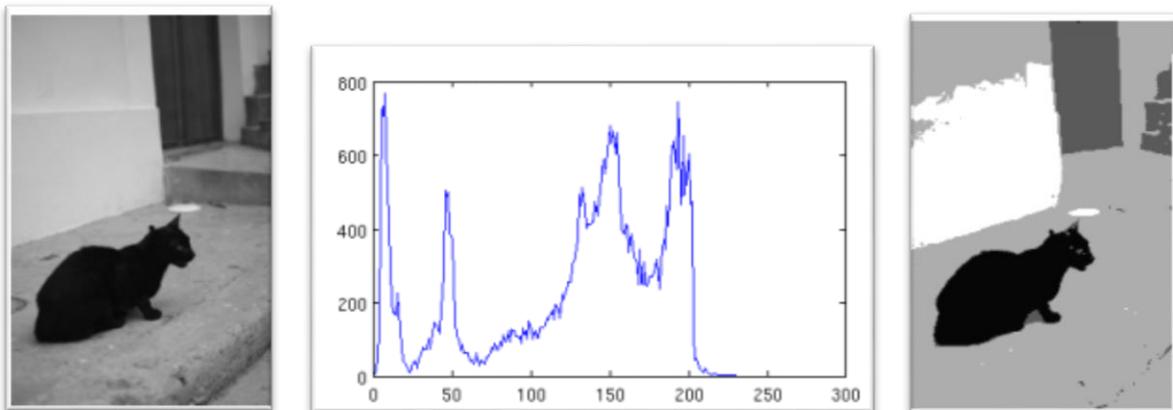


Figure 28 : Segmentation d'images par seuillage (<http://www.univ-orleans.fr>)

de à gauche à droite : image originale ; histogramme de l'image original ; image segment par les seuils 24, 67,182.

III.2 La méthode de segmentation par croissance de régions :

Leur principe repose sur un choix initial d'un ensemble de pixels pour les considérer comme des noyaux (germes), ainsi partir des germes, la croissance des régions s'effectuera en ajoutant de nouveaux pixels qui respectent certain critère d'homogénéité. Comme il est

possible d'avoir la création de nouveaux noyaux dans les régions qui ne pouvant pas être fusionner avec les régions existantes [41].

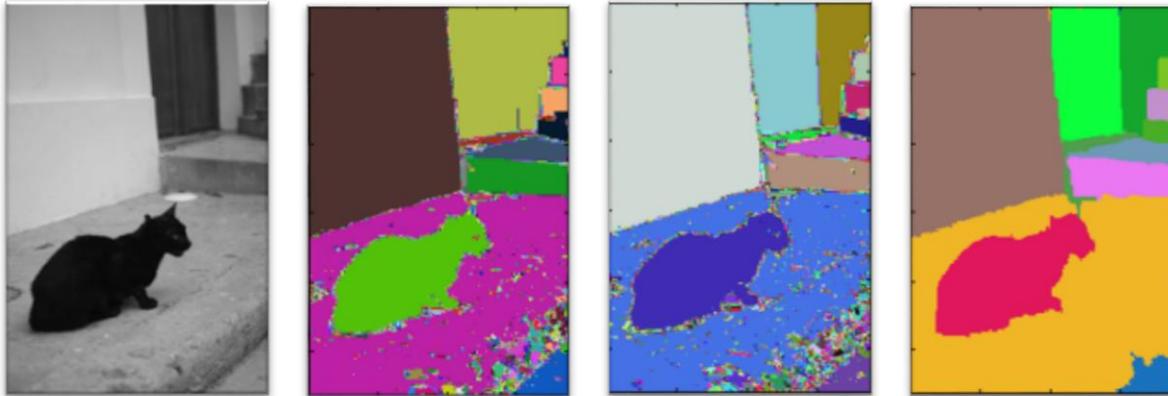


Figure 29 : segmentation par croissance de régions

de à gauche à droite : image originale ; Image segmentée avec $s = 5$;

image segmentée avec $s = 4$; Image segmentée avec $s = 4$, $\text{sup}=200$

(sup signifie que les régions de moins de pixels ont été supprimées).

III.3 La méthode de segmentation par les contours :

Le principe de la détection de contours repose donc sur l'étude des dérivées de la fonction d'intensité dans l'image

La détection de contour dans une image est une étape préliminaire dans de nombreuses applications de l'analyse d'images. Les contours constituent des indices riches pour toute interprétation ultérieure.

Les contours dans une image proviennent des discontinuités de la fonction d'intensité dans les images en raison d'une texture, d'un ombre, des bords de l'objet, etc [44].

Dans la segmentation par contours, il y a des deux problématiques à savoir et à résoudre :

III.3.1 La détection de contours :

La détection de contours est une étape préliminaire à de nombreuses applications de l'analyse d'images. Les contours constituent des indices riches, au même titre que les points d'intérêts, pour toute interprétation de l'image.

Les contours dans une image proviennent des :

- Discontinuités de la fonction de réflectance (texture, ombre) ;
- Discontinuités de profondeur (bords de l'objet).

Pour les détecter, il existe deux types d'approches :

- Approches gradient : détermination de s extrema locaux dans la direction du gradient.

- Approche laplacien : détermination des passages par zéro du laplacien [45].

III.3.2 Fermeture des contours :

En général, après avoir déterminés les contours de l'image, il est nécessaire de chaîner les pixels contours afin qu'ils constituent des lignes fermées correspondant aux frontières des régions.

L'idée est donc de prolonger les segments ouverts afin qu'ils deviennent connexes et qu'ils permettent une partition en régions de l'image segmentée.

Cette prolongation de segments se fait par apport au voisinage de chaque pixel et doit donner, on l'espère, une fermeture de contours cohérente.

Il est à noter que les méthodes de détection des pixels contours sont particulièrement bien adaptées à la segmentation d'images où les couleurs sont fortement contrastées.

Les résultats sont alors parfois décevants car les contours dans l'image peuvent être difficiles à percevoir, ce qui implique l'oubli de certaines régions [45].

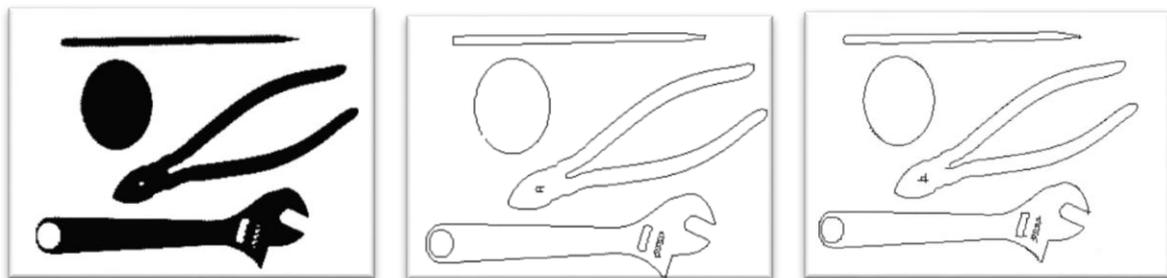


Figure 30 : segmentation par les contours (<https://perso.telecom-paristech.fr>)

De à gauche à droite : image originale ; image segmenté par gradient ;
image segmente par laplacien

III.4 Les méthodes de décomposition/ fusion (split and merge) :

L'idée de cette méthode qui se divise en deux phases, le split (décomposition) et le merge (fusion), est de regrouper les pixels de l'image originale en zones homogènes pré-calculées sur l'image. De ce fait elle adopte le principe de l'approche hybride [46].

La première, split, consiste à diviser (décomposition) l'image en plusieurs régions. On part de l'image de départ. On fait un test d'homogénéité. Si on considère la région homogène on s'arrête, sinon on divise cette région en 4 régions. On continue ainsi sur chaque sous

régions et on s'arrête lorsque toutes ces régions sont homogènes, la plus petite région possible étant le pixel, qui est forcément homogène.

Si on considère l'image I :

$$I = \begin{bmatrix} 10 & 10 & 10 & 40 & 40 & 60 & 60 & 200 \\ 10 & 10 & 10 & 40 & 40 & 60 & 60 & 200 \\ 80 & 100 & 100 & 40 & 40 & 60 & 60 & 200 \\ 80 & 100 & 100 & 150 & 150 & 60 & 60 & 200 \\ 80 & 1 & 1 & 150 & 150 & 60 & 60 & 200 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 200 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 200 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 200 \end{bmatrix}$$

Avec un critère d'homogénéité tel que tous les pixels d'une région aient la même intensité l'étape split donne le résultat suivant (**Fig 31**) divisant ainsi l'image I en 52 régions:

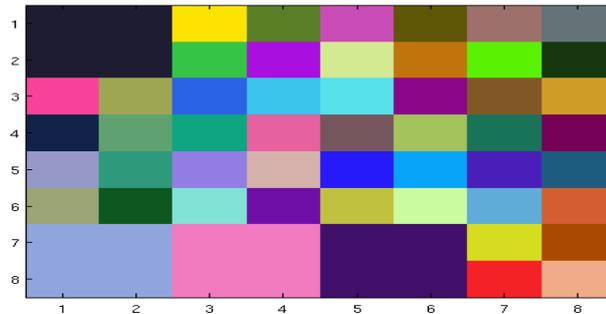


Figure 31 : Split de l'image I [46].

Les régions étant découpées en 4 à chaque étape, des zones homogènes de l'image peuvent alors être séparées. C'est pourquoi, l'étape merge (fusion) est chargée de regrouper les régions voisines si elles vérifient le critère d'homogénéité. C'est-à-dire, pour l'image I :

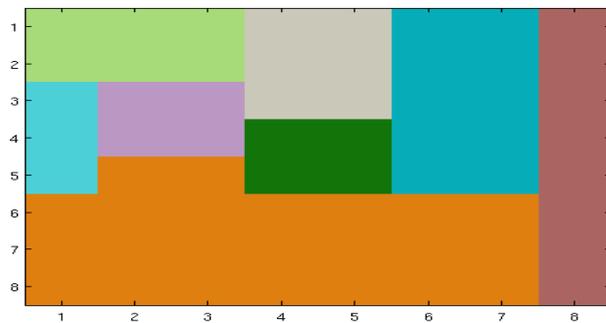


Figure 32: Merge de l'image I [46].

III.5 Segmentation avec Graph cut (coupe de graph) :

La dernière méthode de segmentation et la coupe de graphe (*graph cut* en anglais), nous donnons en premier une définition et des notions de base sur les graphes [47]. L'objectif est de segmenter les objets principaux d'une image en utilisant cette méthode basée sur des coupes de graphes (*graph cut*) [48].

III.5.1 Définition :

Un graphe est un schéma qui permet de décrire un ensemble d'objets et leurs relations, c'est-à-dire le lien entre les objets.

Les objets sont appelés les nœuds ou les sommets du graphe, un lien entre deux objets est appelé une arête.

Un graphe $G = (V, E)$ est défini par l'ensemble $V = \{v_1, v_2, \dots, v_n\}$ non vide dont les éléments sont appelés sommets, ou bien nœuds (*Vertices* en anglais), et par l'ensemble $E = \{e_1, e_2, \dots, e_m\}$ dont les éléments sont appelés arêtes, ou bien arcs (*Edges* en anglais) [47].

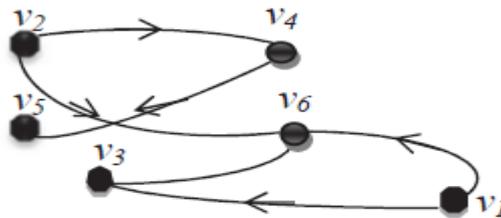


Figure 33: Exemple d'un graphe [47].

III.5.2 Graph cut (coupe de graph) :

Une coupe de graphe valué $G = (V, E, w)$ possédant un seul sommet S et un seul sommet T , une partition des sommets en deux sous-ensemble disjoints notée (A, \bar{A}) telle que [47]:

$$V = A \cup \bar{A}$$

$$A \cap \bar{A} = \emptyset$$

$$S \in A \text{ et } T \in \bar{A}$$

Où : A est l'objet, \bar{A} est l'arrière plan, et w est la capacité de la coupe c-à-d égal à la somme des arcs qui relient le sous-ensemble A avec le sous-ensemble \bar{A}

$$w(A, \bar{A}) = 5+1+3 = 9$$

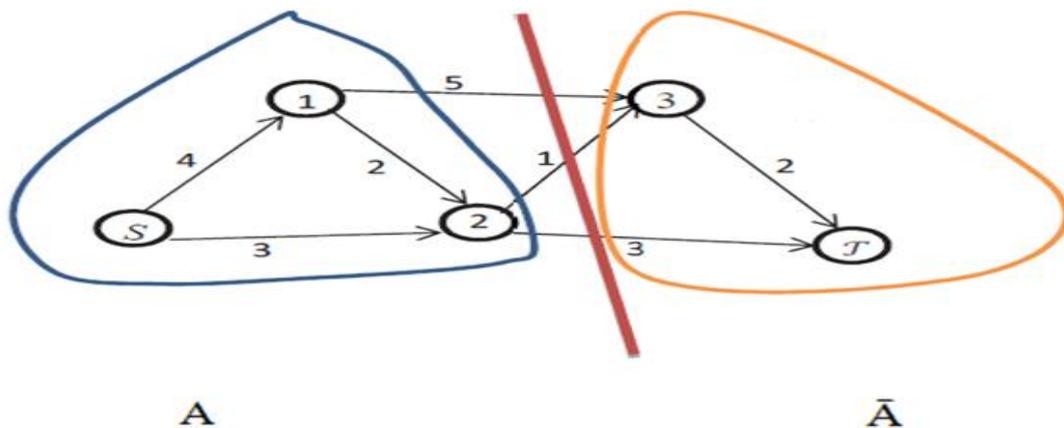


Figure 34: Exemple d'une coupe de graphe [47].

III.5.3 Théorème de MAX FLOW/ MIN CUT :

La coupe minimale est la coupe qui a la minimale valeur absolue appelée (*min cut*), et elle peut être obtenue en trouvant le flot maximum (*max flow*). Le *min cut* est l'équivalent du *max flow* [49]. L'algorithme de *min cut/ max flow* a été développé en 2004 par Boykov and Kolmogorov [50], et peut être utilisé pour obtenir la coupe minimale pour le graphe S-T.

Donc le graphe est divisé par cette découpe, les nœuds sont séparés en deux sous-ensembles disjoints S et T .

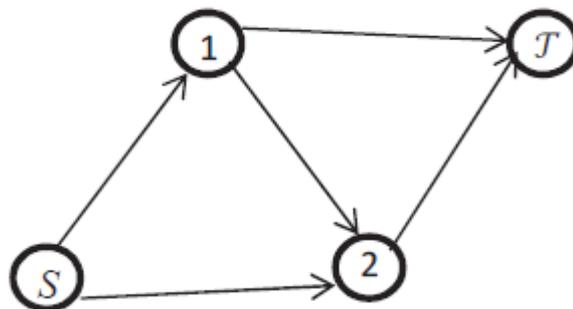


Figure 35: Exemple d'un graph S-T [47].

Où : S représente l'objet contient que des arcs sortant, et T représente l'arrière plan contient que des arcs entrant.

La segmentation d'image par la théorie des graphes est une méthode basée sur l'énergie. *Graph cuts* elle se traduit par :

$$|C| = \sum_{e \in C} w_e \quad 4.5$$

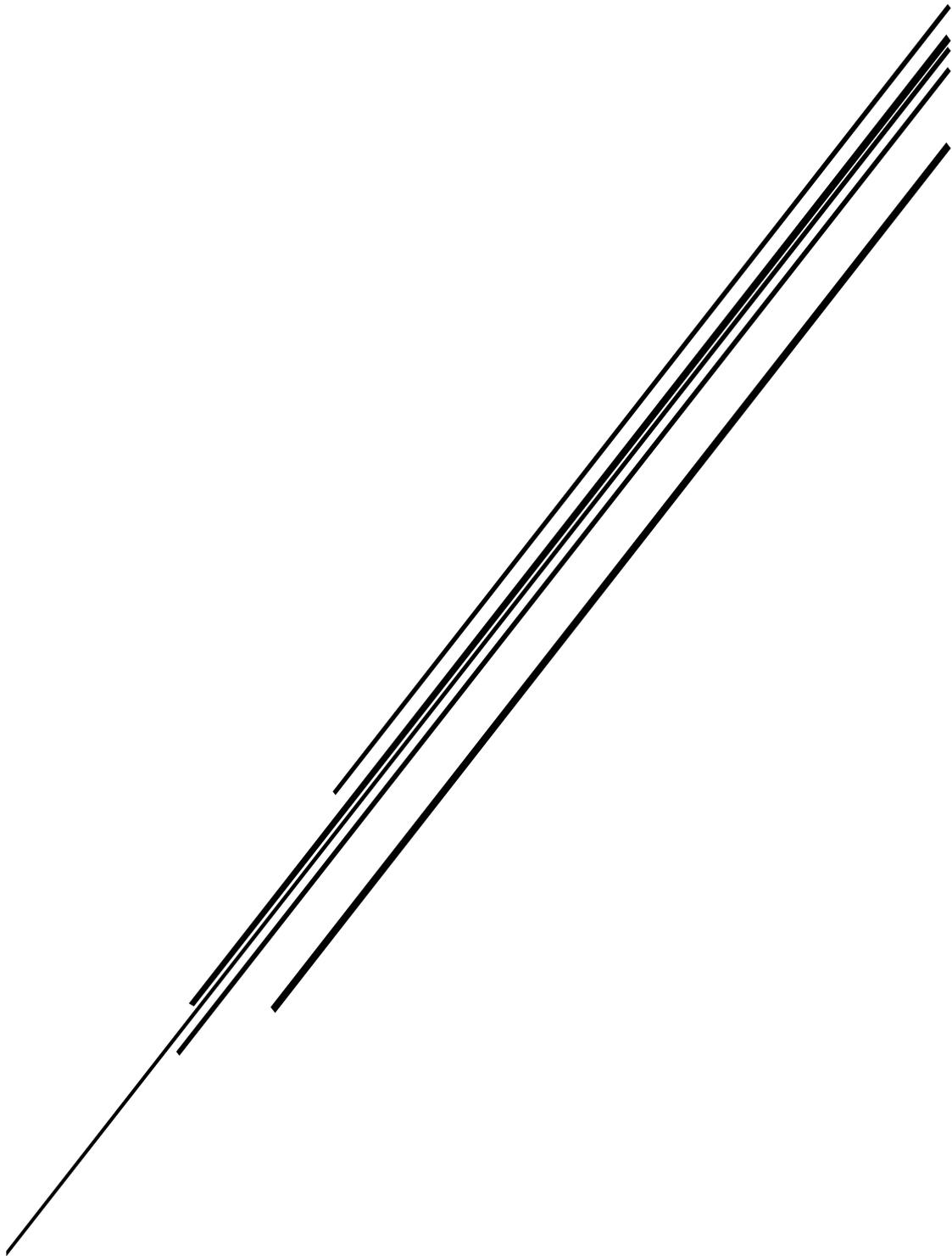
Où : w est la capacité de la coupe, C est le coût d'une coupe de graph, e est l'ensemble des arrêts.

IV Conclusion :

Le choix de la méthode de segmentation est fortement lié non seulement à la nature de l'application visée (détection d'une ou de quelques régions aux caractéristiques connues a priori de l'ensemble de l'image, détection de contours ...etc.) mais aussi à la nature de l'image à traiter à savoir la qualité de ces contours, présence de bruit ou non...etc.

De ce fait, il est impératif de prendre en considération toute les informations possibles en compte avant de trancher sur la méthode à utiliser afin de garantir un fonctionnement performant du système de segmentation de l'image.

Chapitre 5 : Choix de la méthode & mise en œuvre



I Introduction :

Après avoir présenté dans les chapitres précédents les images numériques et le traitement que nous pouvons effectuer sur ces images ainsi nous avons vu la saillance pour détecter les objets, enfin les méthodes de la segmentation de ces objets dans une image.

Dans ce présent chapitre nous traitons les différentes étapes de l'analyse, ainsi l'application des algorithmes sur les images. L'algorithme chaîne markovienne pour détecter l'objet saillant dans l'image, ensuite `graph_cut` pour segmenter cet objet. Ainsi nous structurons ce chapitre comme suit : dans une première section nous présentons l'environnement du travail, dans la seconde section nous décrivons les méthodes et l'analyse des résultats obtenue à travers quelques images résultantes

Dans cette partie nous présentons l'environnement logiciel où nous avons eu les résultats souhaitent.

II Environnement du travail :

Le choix des outils de programmation se fait par plusieurs facteurs : La puissance, La disponibilité de plusieurs fonctionnalités, ...etc. Dans le cadre de notre projet, nous avons utilisé l'outil suivant : MATLAB R2014a

- **MATLAB :**

MATLAB est un langage de programmation simple signifie : « **Matrix laboratory** », rapide, universel de quatrième génération développée par la société The MathWorks. Il a été initialement conçu par Cleve Moler à la fin des années 1970.

MATLAB est un environnement puissant, complet en plus de sa disponibilité est assurée sur plusieurs plateformes : Sun, Bull, HP, IBM, compatibles PC (DOS, Unix ou Windows), Macintosh, iMac et Plusieurs machines parallèles. Matlab est aussi un environnement de calcul scientifique et de visualisation de données qui existe depuis plus de dix ans. Sa facilité d'apprentissage et d'utilisation (due à une syntaxe très claire). Il est utilisé pour effectuer du calcul numérique, manipuler des matrices, afficher des courbes et des données, ... [51].

Parmi les raisons qui nous ont poussés à l'utiliser, on trouve :

- Ses très nombreuses fonctions prédéfinies et prêtes à l'emploi.
- Sa simplicité à l'implémentation et rapidité de calculs.
- Sa fiabilité et sa robustesse.

La figure suivante illustre l'interface d'accueil de MATLAB R2014a :

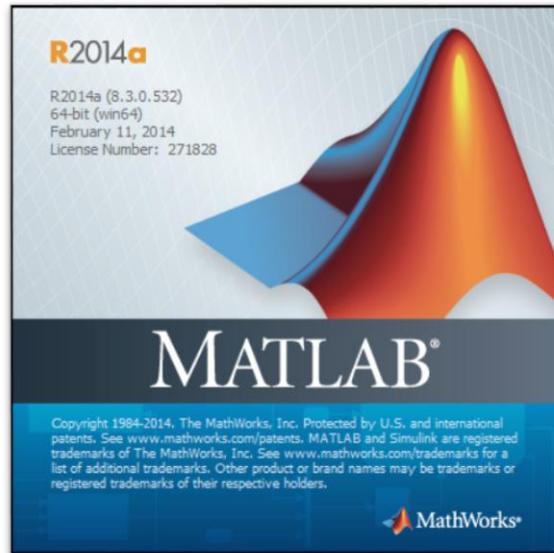


Figure 36: La page d'accueil de MATLAB R2014a

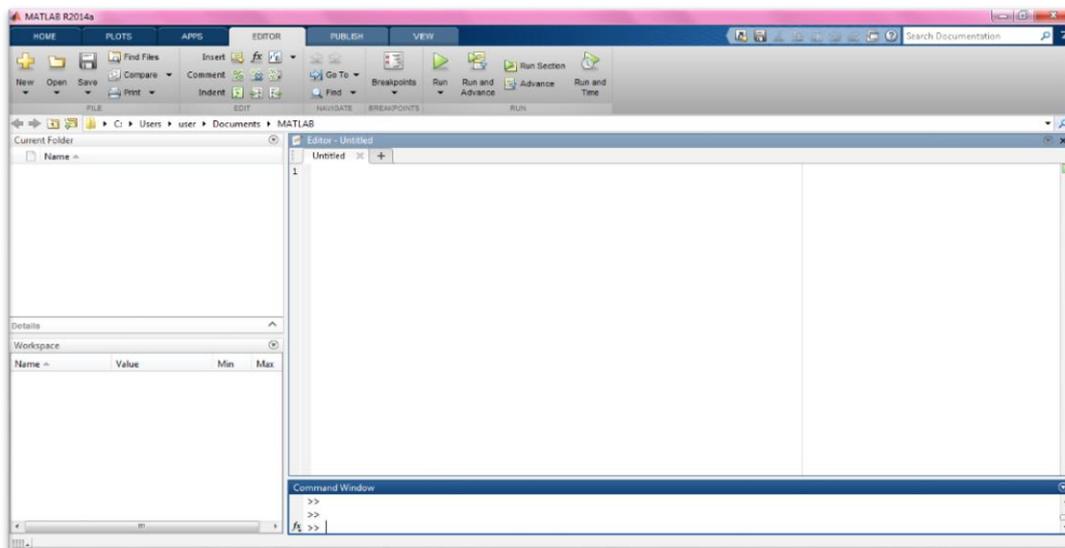


Figure 37 : fenêtre principal de matlab

III La détection de la saillance avec Chaîne de Markov absorbante :

La détection de la saillance via une chaîne de Markov est l'un des méthodes basées sur les graphes ont émergé comme un outil efficace pour la détection d'objets saillants. Ces méthodes permettent de représenter l'image sous forme d'un graphe dont les nœuds sont des régions de l'image et les arcs des liens pondérés reliant les différents nœuds. En plus de leur simplicité, ces méthodes permettent d'incorporer facilement les informations à priori sur la

position spatiale comme la contiguïté des parties de l'objet saillant ainsi que du fond de l'image et l'information à priori sur l'arrière-plan.

Cette méthode permet de segmenter une image en K régions appelées régions (superpixels) en fonction du contenu de l'image. La segmentation en superpixels est très utile pour faire de la détection d'objets, pour découper des images et pour alléger la quantité de données à traiter.

L'algorithme peut se résumer avec les étapes suivantes : [52]

1) D'abord segmenter l'image en K régions (superpixels) ;

2) construire un graphe $G(V, E)$ où les nœuds V sont représentés par les superpixels et les arcs E des liens pondérés reliant les nœuds ;

- Notons que les arcs relient uniquement chaque nœud à ses voisins transitoires. Les nœuds absorbants ne sont donc pas interconnectés ;

- Le poids reliant deux nœud i et j est défini comme suit :

$$W_{i,j} = e^{-\frac{\|x_i - x_j\|}{\sigma^2}} \text{ avec } i, j \in V \quad 5.1$$

➤ Où x_i et x_j les moyennes des nœuds i et j respectivement dans l'espace de couleur CIE $L^*a^*b^*$ (La couleur $L^*a^*b^*$ est constituée d'une composante de luminance ou luminosité L^* et de deux composantes chromatiques : la composante a^* (de vert à rouge) et la composante b^* (de bleu à jaune).) ;

➤ σ une constante qui contrôle la stabilité du poids.

- Réorganiser Les nœuds de façon que les t premiers nœuds soient transitoires et les r derniers soient absorbants.

3) on calcule les matrices suivantes :

- On définit alors la matrice d'affinité A comme suit :

$$A: a_{i,j} \begin{cases} w_{i,j} & j \in N(i), 1 \leq i \leq t \\ 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases} \quad 5.2$$

Avec $N(i)$ l'ensemble des nœuds connectés à i ;

- La matrice de transition P est alors calculée comme suit :

$$P = D^{-1} \times A \quad 5.3$$

➤ avec D la matrice des degrés qui représente la somme des poids connectés à chaque nœud

$$D = \text{diag}(\sum_j a_{ij}), j \in t + r. \quad 5.4$$

➤ Comme le graphe ne relie un nœud qu'à ses voisins, la matrice P est donc éparsée. Cela restreint les marches aléatoires à un seul mouvement pour chaque étape dans le graphe.

4) On extrait la matrice Q à partir de p et on calcule la matrice N

- La matrice P possède la forme canonique suivante :

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix} \quad 5.5$$

- avec $Q \in [0,1]^{t \times t}$ la matrice contenant les probabilités de transition entre chaque paire de nœuds transitoires ;
- $R \in [0,1]^{t \times r}$ la matrice contenant les probabilités de transition entre chaque état transitoire et chaque état absorbant ;
- La matrice 0 représente une matrice de 0 de taille $r \times t$;
- I la matrice identité de taille $r \times r$.

- La matrice fondamentale N est définie comme suit : $N = (I_{t \times t} - Q)^{-1}$ où l'élément n_{ij} présente le nombre de transitions sur le nœud j sachant que la chaîne a commencé à partir de i .

➤ La somme $\sum_j n_{ij}$ reflète le nombre de transitions nécessaire avant l'absorption sachant que la chaîne commence à partir du nœud i . La normalisation de ce nombre dans l'intervalle $[0,1]$ correspond au niveau de saillance du nœud i . Autrement dit, plus il y a de transitions entre i et des nœuds transitoires avant d'atteindre un nœud absorbant, plus il est différent de l'arrière-plan et plus il est saillant.

- On calcule le temps absorbé pour chaque état transitoire comme suit :

$$\bar{Y} = N \times c \quad 5.6$$

- où c est un vecteur de taille t dont les éléments sont tous égaux à 1
- \bar{y} une valeur normalisée dans $[0,1]$.

- On obtient une carte de saillance S avec :

$$S(i) = \bar{y}(i) \quad i = 1, 2, \dots, t, \quad 5.7$$

- où i indexe les nœuds transitoires sur le graphe ;
- \bar{y} indique le vecteur de temps absorbé normalisé.

Etant donné que le nombre de transitions à partir du centre de l'image jusqu'à ses bords peut être élevé, les parties de qui se situent au centre de l'image sont plus susceptibles d'être considérées comme saillantes. D'une part, cette affirmation n'est pas toujours correcte et cela

peut sensiblement générer un bruit sur l'image surtout dans le cas de parties d'arrière-plan à surfaces larges et à caractère homogène. D'une autre part, une mise à jour de l'estimation de la saillance de telle manière à réduire les valeurs de saillance au centre de l'image peut affecter le taux de vrais positifs et considérer ainsi des parties saillantes comme arrière-plan. On se base donc sur l'observation qu'une bonne carte de saillance est caractérisée par un contraste élevé entre l'objet saillant et l'arrière-plan.

- Pour pallier ce problème, nous mettons à jour la carte de saillance en utilisant un temps absorbé pondéré \mathbf{y}_w , qui peut être noté comme suit :

$$\mathbf{y}_w = N \times \mathbf{u} \quad 5.8$$

- où \mathbf{u} est le vecteur de colonne de pondération. Dans ce travail, nous utilisons le temps récurrent normalisé d'une chaîne de Markov ergodique, dont la matrice de transition est la ligne normalisée Q comme le poids u .

5) on calcule un score pour estimer le degré de ce contraste comme suit :

$$\text{score} = \sum_{b=1}^{n10} g(b) \times \min(b, (11 - b)) \quad 5.9$$

- avec g l'histogramme du niveau de gris de la carte de saillance
- et $b \in \{1, 2, \dots, 11\}$ représente un bin de l'histogramme.
- Si la valeur du **score** est supérieure à un certain seuil γ , alors une mise à jour de l'estimation de la saillance est requise, autrement aucune mise à jour n'est nécessaire.

6) On calcule le temps récurrent h , et le poids u , puis on calcule la carte de saillance S

- La mise à jour s'opère en considérant une chaîne de Markov ergodique. Soit π la distribution d'équilibre pour la chaîne de Markov ergodique considérée. Celle-ci est calculée à partir de la matrice d'affinité A comme suit :

$$\pi_i = \frac{\sum_j a_{ij}}{\sum_{ij} a_{ij}} \quad 5.10$$

- avec i et j les indexes des nœuds transitoires.
 - La valeur de la somme $\sum_j a_{ij}$ est particulièrement élevée dans les zones homogènes de l'image.
 - Le temps récurrent dans ce type de régions est donc faible. Pour cette raison, on utilise le temps récurrent moyen :

$$\mathbf{h}_i = \mathbf{1}/\pi_i \quad 5.11$$

- i indexe tous les états de la chaîne de Markov ergodique ;
- \mathbf{h}_i le temps de récurrence moyen ;

➤ π_i décrit la probabilité attendue ;

D'où l'état d'équilibre est caractérisé par la distribution d'équilibre π , qui satisfait à l'équation

$$\pi P = \pi \quad 5.12$$

Avec π est un vecteur de probabilité strictement positif, et P est la matrice de transition ergodique

- Ils substituent ensuite le vecteur u défini comme suit :

$$u_j = h_j / \sum_k h_k \quad 5.13$$

➤ avec k et j des indexes des nœuds transitoires sur le graphe.

Puis on calculera saillance S avec l'eqV.8 Et l'eqV.7. Enfin on aura comme une sortie une carte de la saillance.

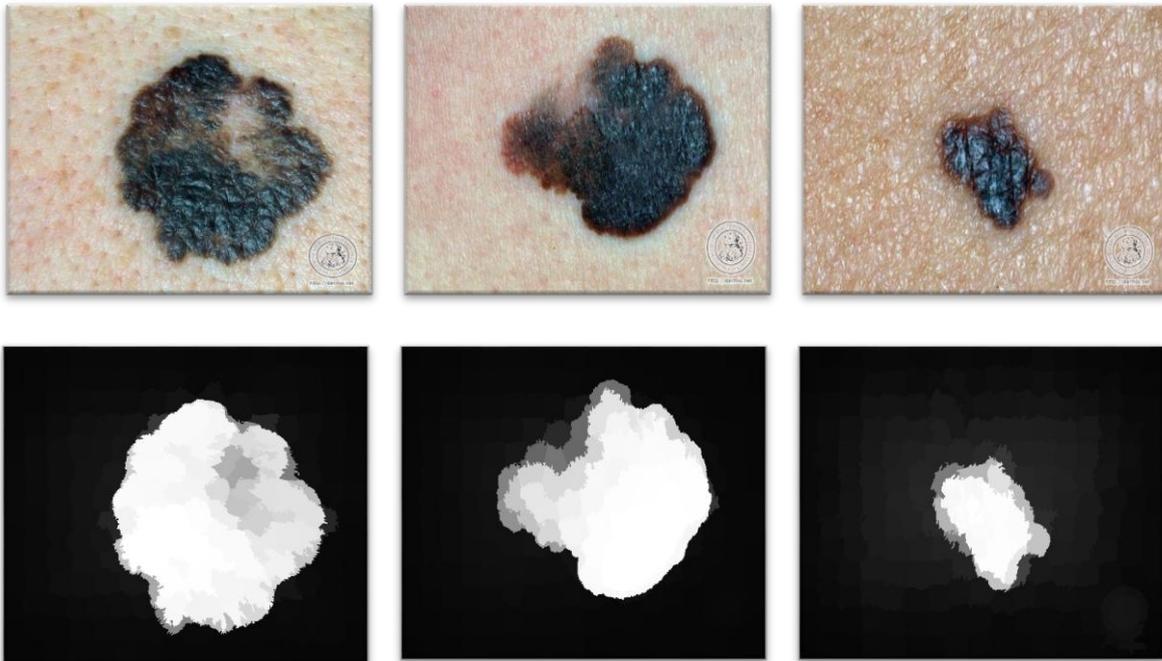


Figure 38 : les images générées avec « chaine de markov »

En dessus images originales ; en dessous les cartes de la saillance avec paramètre $k=250$.

IV La segmentation avec « graph_cut » :

Après avoir la carte de la saillance nous allons faire la segmentation avec graph_cut (la coupe de graphe), L'objectif du graph_cut est de diviser les nœuds du graphe en deux groupes : le groupe S qui sont les nœuds de départ et le groupe T qui sont les nœuds d'arrivée, et d'attribuer à chaque pixel une classe w pouvant être soit «O» si le pixel considéré appartient à l'objet, soit «B» s'il appartient au fond.

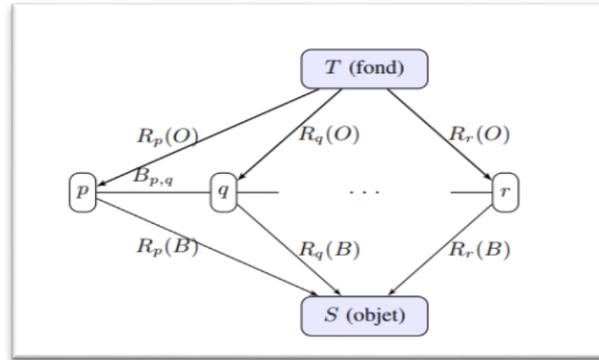


Figure 39 : Exemple d'un graphe et les liens pondérés, avec p, q, r des nœuds du graphe et S et T les nœuds terminaux.

Nous exposons ici la méthode de segmentation binaire d'images par coupure dans un graphe : [53]

- On définit un graphe $G = \langle V, E \rangle$, où V est l'ensemble des nœuds et E l'ensemble des segments.

- L'image I à segmenter est considérée comme étant ce graphe, chaque pixel en étant un nœud.

- La segmentation de l'image est exprimée sous la forme d'un tableau de valeurs d'opacité $\alpha = (\alpha_1, \dots, \alpha_N)$ à chaque pixel. Généralement $0 \leq \alpha_n \leq 1$, mais pour la segmentation dure $\alpha_n \in \{0,1\}$, avec 0 pour le fond et 1 pour le premier plan. Les paramètres θ décrivent les distributions du niveau de gris et du premier plan de l'image et consistent en des histogrammes de valeurs de gris : $\theta = \{h(z; \alpha), \alpha = 0,1\}$, un pour le fond et un pour le premier plan. Les histogrammes sont assemblés directement à partir de pixels étiquetés provenant des régions trimap respectives TB, TF . (Les histogrammes sont normalisés pour totaliser 1 sur la plage de niveaux de gris : $\int_z h(z; \alpha) = 1$) La tâche de segmentation consiste à déduire les variables d'opacité inconnues α des données d'image données z et du modèle θ .

- Attribuer, à chaque pixel, une classe ω , pouvant être soit O si le pixel considéré appartient à l'objet, soit B s'il appartient au fond.

- On considère deux nœuds supplémentaires, appelés nœuds terminaux : la source S représentant l'objet O, et le puits T représentant le fond B.

- Chaque couple de nœuds $(p, q) \in V^2$ dans un voisinage N est connecté par un segment appelé n-link.

- Chaque nœud $p \in V$ est connecté aux nœuds terminaux S et T par deux segments respectifs appelés t-links (Figure 30). Les liens t-links et n-links sont pondérés par des coûts

définis ci-après. La coupe d'un lien t-link de p vers T (resp. S) permet d'attribuer l'étiquette O (resp. B) à ce pixel, définissant ainsi une segmentation de l'image.

- Le coût d'une coupe C est donné par la somme des poids des segments n-links et t-links concernés. La meilleure segmentation est obtenue en calculant la coupe de coût minimum.

- Soit I_p le niveau de gris du pixel p. Les t-links entre un pixel p et les deux nœuds terminaux S et T sont pondérés par un terme dit terme région, noté R_p et défini par :

$$R_p(\omega) = -\ln P_r(I_p|\omega) \quad 5.14$$

où $(I_p|\omega)$ est la probabilité que le pixel p appartienne à la classe ω , généralement obtenue par un histogramme sur les niveaux de gris de l'objet et du fond.

Les n-links sont pondérés par un terme de régularisation, ou (énergie) contour, conçu pour assurer la cohérence spatiale dans un voisinage de pixels. Ce terme noté $B_{p,q}$ et est classiquement défini par :

$$B_{p,q} \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\text{dist}(p,q)} \quad 5.15$$

où

- I_p et I_q sont les niveaux de gris des pixels p et q ;
- $\text{dist}(p, q)$ la distance euclidienne entre p et q ;
- σ une constante, généralement liée au bruit d'acquisition ;
- Dans une région uniforme, I_p et I_q sont similaires et on a alors

$|I_p - I_q| < \sigma$ Ceci implique une forte valeur pour $B_{p,q}$, décourageant toute coupe du graphe dans ce voisinage.

$|I_p - I_q| > \sigma$ Ceci implique une faible valeur Alors I_p et I_q sont différents et dans une zone de contours.

- L'énergie totale d'une coupe C dans le graphe est alors définie par :

$$E(C) = \sum_{p \in V} R_p(\omega_p) + \lambda \sum_{p,q \in N} B_{p,q} \cdot \delta(\omega_p \neq \omega_q) \quad 5.16$$

où $\delta(\omega_p \neq \omega_q)$: vaut 0 si p et q ont la même étiquette ;

$\delta(\omega_p \neq \omega_q)$: vaut 1 si p et q ont étiquette différente.

- Par la suite, on applique l'algorithme du flot max sur le graphe de l'image afin d'obtenir la coupe minimale. Nous utilisons l'algorithme du flot max/coupe min développé par Yuki Boykov et Vladimir Kolmogorov [54]. Ce nouvel algorithme qui améliore les

performances empiriques de l'algorithme de Ford &Fulkerson est basé aussi sur les chaînes augmentantes. Pour plus de précision sur l'algorithme [55]

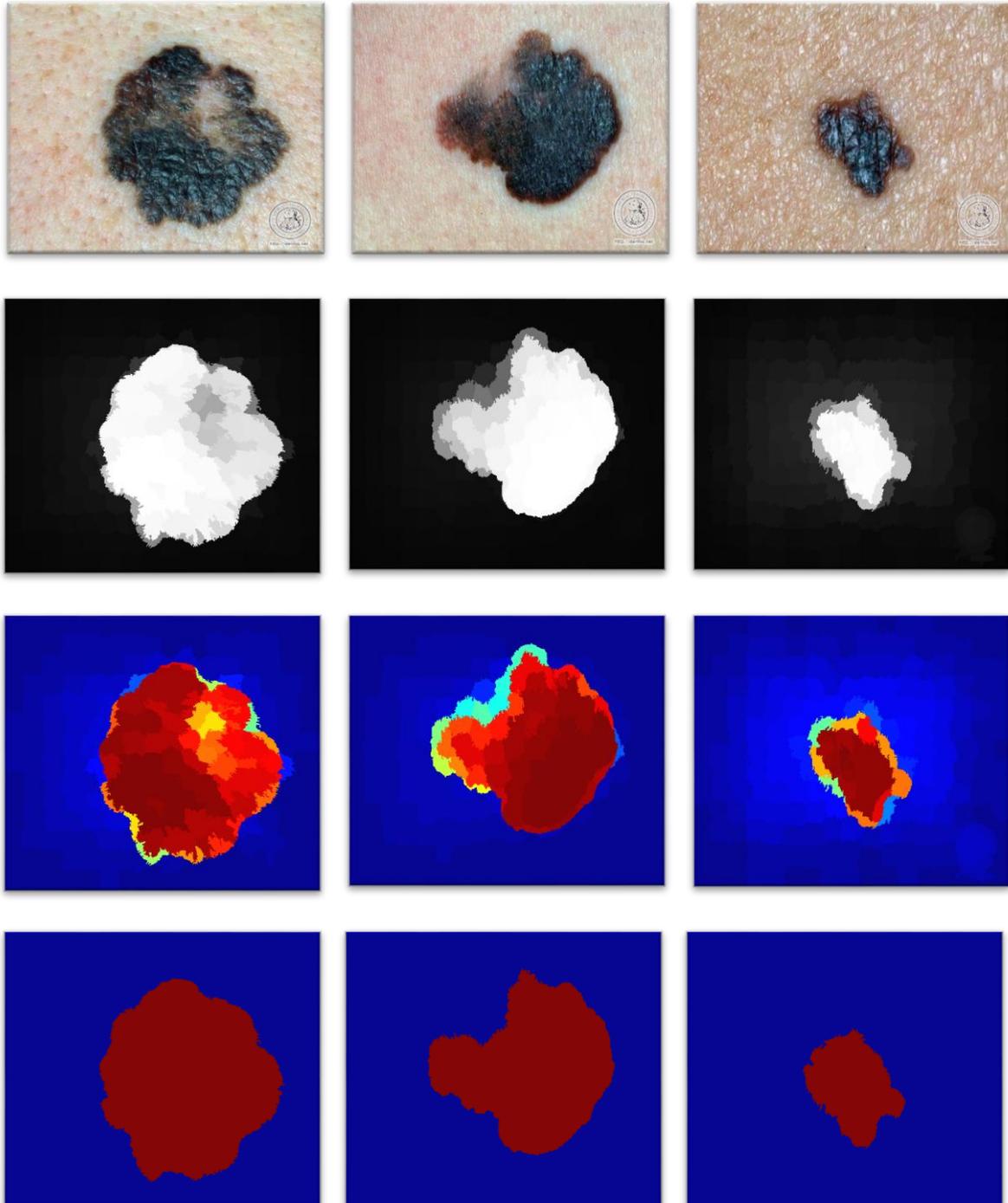


Figure 40 : image segmenter par graph-cut

1^{ère} ligne images originales ; 2^{ème} ligne la carte de saillance ; 3^{ème} ligne l'image probabilités des pixels ; 4^{ème} ligne l'objet et l'arrière-plan avec maxflow /mincut

V Evaluation de la méthode proposée :

Nous avons choisis à décrire quelques méthodes de segmentation des images pour faire une comparaison entre notre approche, et ces méthodes, à fin de trouver celle qui est ou qui sont les plus efficaces dans la segmentation des images, pour détecter avec une grande précision.

Nous commençons tout d'abord par une description globale des différentes techniques de segmentation Nous citons parmi celles : la méthode de segmentation

Multithresh (otsu multi niveau) est une fonction de matlab, c.à.d. Seuils d'image à plusieurs niveaux à l'aide de la méthode d'Otsu . Qui a le principe :

- calculer les deux niveaux de seuil ;
- segmenter l'image en trois niveaux (des pixels d'arrière-plan, des pixels clairs et des pixels intermédiaires.) ;
- convertir l'image segmentée en image en couleurs.

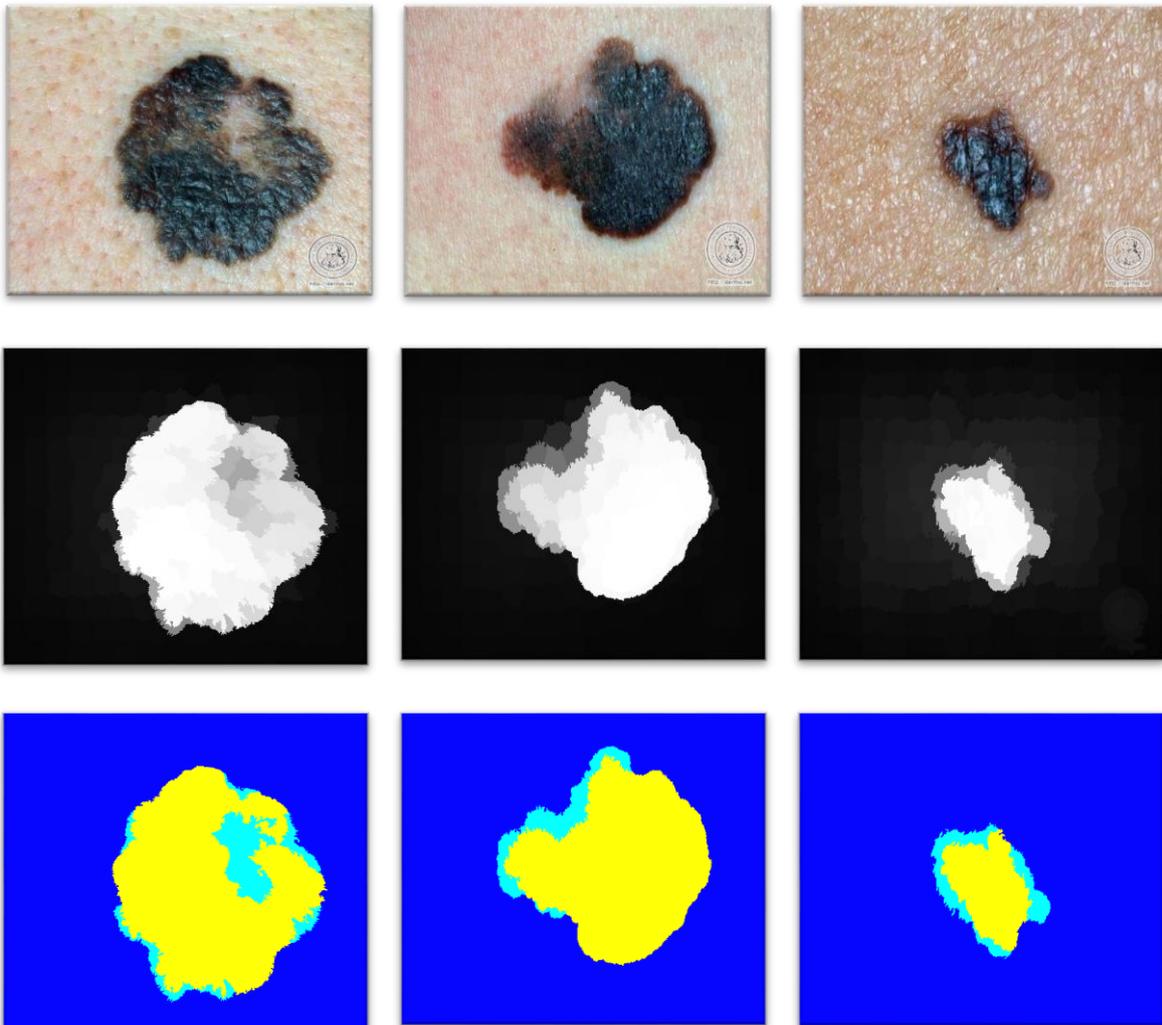




Figure 41 : image segmentée par « otsu » et « multithresh »

1^{ère} ligne les images originales ; 2^{ème} ligne les cartes de saillance associées ; 3^{ème} ligne les images multi_otsu « multithresh » ; 4^{ème} ligne les images segmentées avec « otsu ».

VI Résultats de segmentation par les trois méthodes :

La segmentation par otsu est une méthode qui partitionne une image en deux classes l'objet et l'arrière-plan, avec le calcul de seuil à partir de l'histogramme d'une image. Otsu est simple et facile à utiliser et nous obtenons des résultats rapides, mais comme nous voyons que cette méthode de segmentation donne que la forme de l'objet, pas les caractéristiques de l'objet segmenté.

La segmentation par otsu multi-niveau, nous remarquons que elle a le même principe que otsu, la différence revient à segmenter notre image à plusieurs segments. Une méthode simple à implémenter et donne des résultats rapidement.

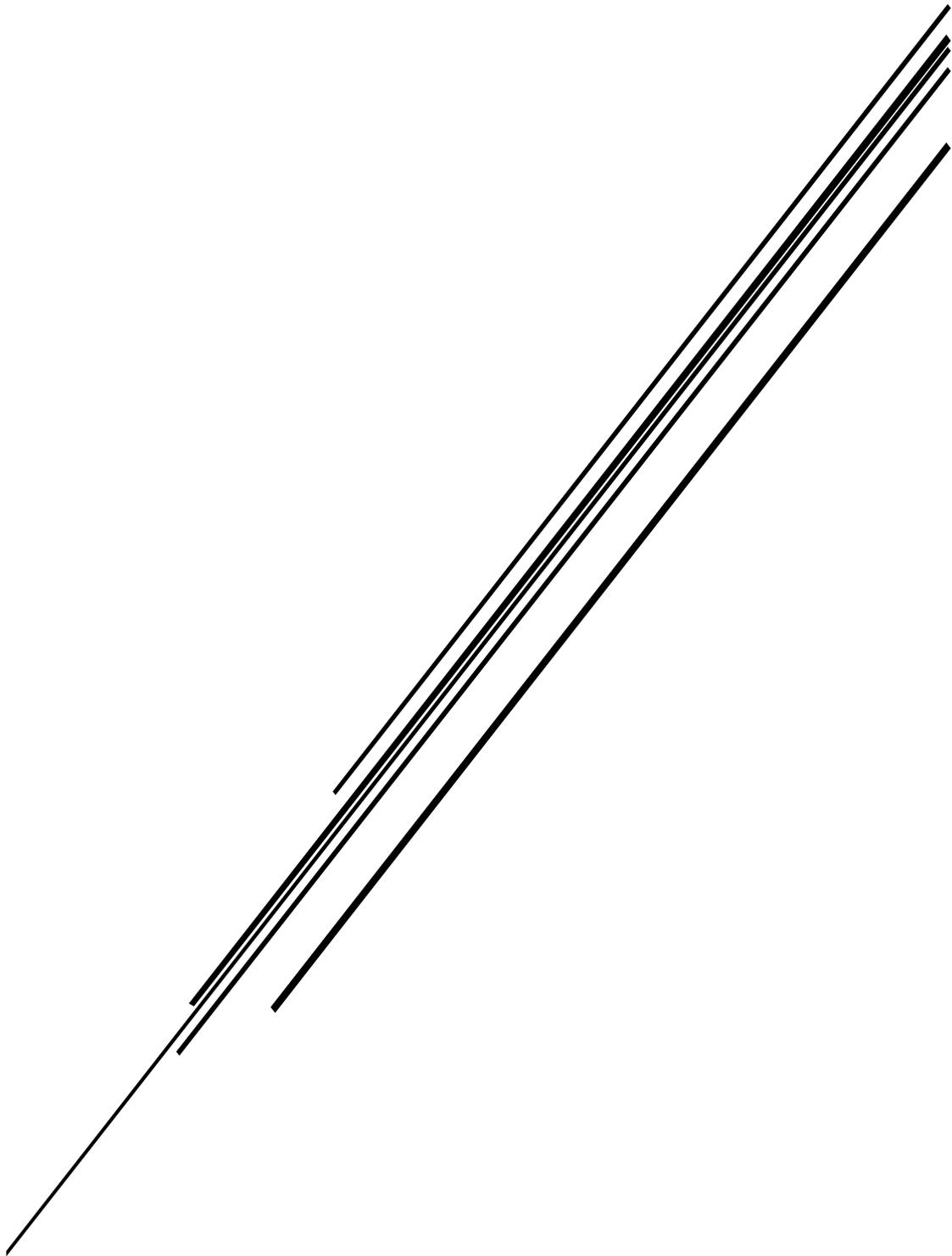
L'algorithme de segmentation Graph-cut est très utile dans la segmentation des images, il donne des bons résultats dans l'extraction exacte de l'objet. Cet algorithme préserve les caractéristiques initiales de l'objet (la taille, la position, la forme, l'intensité des pixels, le contour...). Mais difficile à implémenter et avoir le résultat final.

VII Conclusion :

La détection de l'objet saillant dans l'image en utilisant l'algorithme de CHAÎNE MARKOV et l'algorithme de GRAPH_CUT est une approche performante dans la détection des objets, elle est rapide, efficace

Afin d'évaluer notre travail, nous avons proposé deux autres méthodes de segmentation par OTSU et OTSU MULTI NIVEAU, pour faire une comparaison entre ces algorithmes dans la détection exacte et précise des objets.

CONCLUSION GENERALE



Dans ce mémoire, nous nous sommes intéressés à la détection et la segmentation des objets saillants sur les images.

Nous avons en premier lieu porté notre réflexion sur la représentation des notions générales sur l'image numérique, le traitement et la segmentation des images.

Nous avons aussi proposé une représentation de traitement d'images qui est une discipline de l'informatique et des mathématiques appliquées qui étudie les images numériques et leur transformation dans le but d'améliorer leur qualité ou d'en extraire de l'information.

Nous avons ensuite utilisé la saillance d'objet qui est la manière de modéliser l'attention visuelle en informatique, pour bien expliquer cette notion nous avons vu quelques facteurs physiques et cognitifs de la saillance visuelle et différentes méthodes de cette technique.

En deuxième lieu on a porté notre réflexion sur la segmentation d'objet qui est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis, on a vu aussi quelques méthodes de segmentation telles que la binarisation, la segmentation par les contours, par croissance de région, la méthode (composition/fusion) et le graph cut qui est une méthode très importante dans le domaine du traitement d'images, etc.

En effet, nous avons conçu et réalisé un système de détection d'un objet saillant d'une image. Dans ces modes de détection et de segmentation basés sur une sortie d'une carte de la saillance d'une image puis la segmentée.

D'abord on a utilisé une méthode basée sur les graphes pour détecter l'objet saillant, qui est la chaîne de Markov qui transforme une image en une carte de saillance.

Ensuite on effectue sur cette carte de saillance une approche qui se base sur la théorie des graphes, qui est graph-cut ce principe de segmenter l'objet saillant et l'arrière-plan d'une image, puis on va comparer les résultats de notre approche avec les résultats de la méthode Otsu.

Ce travail effectué nous a donné d'une part de concevoir et de modéliser avec le langage de programmation MATLAB R2014a, et d'autre part d'apprendre et de découvrir l'utilisation de ce langage.

En dernier lieu, nous avons remarqué que la méthode chaîne Markovienne utilisée pour détecter la saillance a montré son efficacité par rapport à d'autres méthodes citées dans le chapitre III. Par ailleurs cette méthode peut avoir de meilleurs résultats chose que nous proposons à faire au futur.

Références bibliographiques :

- [1] : Rapport « reconnaissance en ligne de chiffres manuscrites isolés par réseaux de neurones artificiels » Présenté par : SAFRI Madjid, DEMRI Melissa « 2009/2010 »
- [2] : Rapport « le traitement numérique des images » Présenté par : Peyré, Gabriel « Paris-2011 »
- [3] : Rapport « Binarisation d'images -deux solutions à ce problème- » Présenté par : CHIHKIAN Alian « 1989 »
- [4] : Article « Traitement d'images en JAVA » Présenté par : Slim, Mohamed « 03/07/2010 »
- [5] :Le lien : <http://www.imedias.pro/cours-en-ligne/graphisme-design/definition-resolution-taille-image/les-images-vectorielles-matricielle/>
- [6] : Mémoire de fin d'études pour l'obtention de diplôme de master en informatique. Thème « Segmentation des images par contour actifs : Application sur les images satellitaires à haute résolutions ». Présenté par : MEDJAOUI Amina, FARES Fadia. Option SIC (Système d'Information et de Connaissance)
- [7] : Thèse pour obtenir le grade de docteur en informatique « Reconnaissance et classification d'images de documents » Présenté par : AUGEREAU Olivier « Bordeaux2013»
- [8] : Rapport, Présenté par : BENABDELLAH Mohammed « 2007 »
- [9] : Article « Les extensions de fichiers » « 2008 »
- [10] : Rapport « Qu'est-ce que la luminance? » « 2008 »
- [11] : Rapport « Traitements d'images» « 18/05/2017»
- [12] : Rapport « Algorithme pour le traitement de l'image » Présenté par : Ortner, Mathias « 2004 »
- [13] : Mémoire pour obtenir le diplôme d'ingénieur « Comparaison des différents opérateurs de détection de contours » Présenté par : Feriel BOUDJEMAI « 2008/2009 »
- [14] : Mémoire pour obtenir le diplôme d'ingénieur « Segmentation d'images multi-composantes par coopération des méthodes » Présenté par : OUALI Mohand-Akli, MOHAMDI Kahina « 2008/2009 »
- [15] : Thèse pour obtenir le grade de docteur « Segmentation d'images par combinaison adaptative couleur/texture et classification de pixels » Présenté par : Dhouha ATTIA « Université de Technologie de Belfort-Montbéliard – 03/10/2013 »
- [16] : Rapport « La loi normal » Présenté par : M.Raffestin

- [17] : Rapport « Quelques méthodes mathématiques pour le traitement d'image »
Présenté par : Bergounioux, Maitine « 28/01/2009 »
- [18] : Mémoire pour obtenir le diplôme de master en informatique « Segmentation des images par Contours Actifs : Application sur les Images Satellitaires à Haute Résolutions »
Présenté par : MEDJAOUI Amina, FARES Fadia « 2011/2012 »
- [19] : Mémoire pour obtenir le diplôme de master en informatique, Présenté par : Hacène, BELKACEMI « 2015/2016 »
- [20] : Rapport « Morphologie mathématique » « 2006 »
- [21] : Le lien : <https://dpt-info.u-strasbg.fr/~cronse/TIDOC/MM/deof.html>
- [22] : Le lien : <http://clouard.users.greyc.fr/Pantheon/experiments/illumination-correction>
- [23] : Article « La saillance : questions méthodologiques autour d'une notion multifactorielle » Frédéric Landragin « Laboratoire Lattice – UMR 8094 – CNRS, ENS, Université de Paris-3 »
- [24] : Article « Salient Object Detection : A survey »
Présenté par : Ali Borji, Ming Cheng, Qibin Hou, Huaizu Jiang and Jia Li (Computervision and image understanding).
- [25] : Mémoire pour obtenir le diplôme master en informatique « Détection d'objets saillants dans des images en couleurs », Présenté par : MANSOUR Djamila, FERDJI Hafidha « 2014/2015 »
- [26] : Edmonds P.G.A. « Computational Model of Collaboration on Reference in Direction-Giving Dialogues », Ms, Thesis, University of Toronto. 1993.
- [27] : Article « état de l'art sur la détection d'objets saillants »
- [28] : Y. Hu, D. Rajan, L.T. Chia. Adaptive local context suppression of multiple cues for salient visual attention detection. In: Proceedings of the IEEE International Conference on Multimedia and Expo, ICME, 2005.
- [29] : D. Gao, V. Mahadevan, and N. Vasconcelos. The discriminant center surround hypothesis for bottom-up saliency. In: Advances in Neural Information Processing Systems, 2009.
- [30] : J. Harel, C. Koch and P. Perona. Graph-based visual saliency. In: Advances in Neural Information Processing Systems, 545-552, 2007.
- [31] : Y. Wei, F. Wen, W. Zhu and J. Sun. Geodesic saliency using background priors. European Conf. on Computer Vision, 29-42, 2012.

- [32]: M-M. Cheng, N. J. Mitra, X. Huang, P. H. S. Torr, and S-M. Hu. Global Contrast Based Salient Region Detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 37(3): 569-582, 2015.
- [33]: X. Hou and L. Zhang. Saliency Detection: A Spectral Residual Approach. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007.
- [34]: R. Achanta, S.S. Hemami, F.J. Estrada and S. Susstrunk. Frequency-Tuned Salient Region Detection. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1597 – 1604, 2009.
- [35]: B. Jiang, L. Zhang, H. Lu, C. Yang, and M.-H. Yang. Saliency Detection via Absorbing Markov Chain. *IEEE Int'l Conf. on Computer Vision*, 1665-1672, 2013.
- [36]: W. Zhu, S. Liang, Y. Wei and J. Sun. Saliency Optimization from Robust Background Detection. *IEEE Conf. Computer Vision and Pattern Recognition*, 2814-2821, 2014.
- [37]: P. Wang, J. Wang, G. Zeng, J. Feng, H. Zha, and S. Li, Salient Object Detection for Searched Web Images via Global Saliency. *IEEE Conf. on Computer Vision and Pattern Recognition*, 3194-3201, 2012.
- [38]: Z. Liu, W. Zou and O. Le Meur. Saliency Tree: A Novel Saliency Detection Framework. *IEEE Trans. on Image Processing*, 23(5):1937-1952, 2014.
- [39]: P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5): 898–916, 2011.
- [40]: Q. Yan, L. Xu, J. Shi and J. Jia. Hierarchical Saliency Detection. *IEEE Conf. on Computer Vision and Pattern Recognition*, 1155-1162, 2013.
- [41] : Thèse présentée à l'université d'Orléans pour obtenir le grade de docteur thème : «Caractérisation des objets dans une image en vue d'une aide à l'interprétation et d'une compression adaptée au contenu : application aux images échographiques »
Présenté par : Arnaud CAPRI « discipline : Science et Technologies Industrielles »
- [42] : Mémoire pour l'obtention de diplôme de master «Etude comparative des performances des classifieurs Support Vector Machines (SVM) et Fuzzy C-Means (FCM). Application à la segmentation d'images texturées » Présenté par: Saida MECHHAT , le 27/09/2017 Faculté de Génie Electrique et d'Informatique Département Automatique
Option : Commande des Systèmes
- [43] : Mémoire pour l'obtention de diplôme de master « Segmentation d'image' Application aux documents anciens' » Présenté par : LELORE Thibault « laboratoire des sciences de l'information et des systèmes-Université De Nantes- France / 2007 »

- [44] : Thèse en vue de l'obtention de grade de docteur de l'université de Toulouse délivré par l'université de Toulouse III-Paul Sabatier. Spécialité : Système Embarqué. Soutenu le : 27/09/2013 sous le thème « Détection et suivi d'objet par vision fonder sur segmentation par contour actif basé région » Présenté par : Wassima AIT FARES
- [45] : Mémoire pour obtenir le diplôme master en informatique «Détection d'objets saillants dans des images en couleurs », Présenté par : MANSOUR Djamila, FERDJI Hafidha « 2014/2015»
- [46] : Article « Image Master 2 Automatique » « Projet Image: Méthodes de segmentation » ECHEGUT Romain – 2010
- [47]: Article « Segmentation with Graph Cuts »
- [48] : Thèse pour obtenir le grade de docteur spécialité « Génie Biomédicale » Présentée par : Touria BAAKEK. Soutenu le 25/06/2015 sous le thème « Segmentation et classification intelligente des images cérébrales identification de tumeurs »
- [49]: F.Yi, I.Moon, « Image segmentation; A survey of graph cut methods », International conference and informatics (ICSAI), pp 1936-1941, 2012
- [50]: Y.Boykov, M.P.Jolly « Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images ». International Conference on Computer Vision; Vol 1, PP105-112, 2011 C. Rother, V. Kolmogorov, A. Blake « Interactive foreground extraction using iterated graph cuts». SpecialInterest Group on Graphics and Interactive Techniques, PP309-314, 2004
- [51] : Mémoire de fin d'études de master académique Domaine : Mathématiques et Informatique Filière : Informatique Spécialité : Réseau, Mobilité et Systèmes Embarqués Soutenu le : 28/09/2016 sous le thème « Développement d'un système biométrique pour la reconnaissance de visages, basé sur les ondelettes». Présenté par : Mohamed YOUNSI, Younes MEDJTOUH
- [52]:http://openaccess.thecvf.com/content_iccv_2013/papers/Jiang_Saliency_Detection_via_2013_ICCV_paper.pdf
- [53] : <https://hal.archives-ouvertes.fr/hal-00673396/document>
- [54]: Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/maxflow algorithms for energy minimization in vision,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 9, pp. 1124–1137, 2004
- [55]: An experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision (2004).