



UNIVERSITE MOULOU D MAMMERRI DE TIZI-OUZOU



Faculté de Génie Electrique et d'Informatique
Département d'informatique

Mémoire de fin d'études

En vue de l'obtention du diplôme de Master en Informatique

Option : Ingénierie des Systèmes d'Information
Et Conduit projets informatiques

Thème:

DÉTECTION D'OBJETS SAILLANTS BASÉE SUR
LA DIFFUSION PAR GRAPHE

Présenté par :

BELLAL Amina
IOURZIKENE Karima

Devant le jury composé de :

Président : Mr.SADI Sami

Examineur: Mr.RAMDANI Mohamad.

Promoteur :

Mr.FILLALI Idir

Soutenu le :

Année universitaire 2019/2020

Sommaire

Introduction générale :	1
CHAPITRE I : Traitement d'image et segmentation:	2
I.1.Introduction:	2
I.2.Traitement d'une image	2
I.2.1Définition d'une image :	2
I.2.2Caractéristique d'image :	2
I.3.3 Les types des images :	9
I.3.4Opération de filtrages	10
II.Segmentation :	14
II.1Définition :	14
II.2.Approche contour:	14
II.3.Segmentation basée sur région :	15
II.4.Segmentation par contour actif	16
II.5. Segmentation Superpixel	17
Conclusion	18
Chapitre II : Détection d'objets saillants:	19
II.1.Introduction	19
II.2.Qu'est-ce que la saillance?	19
II.3 .Historique de la détection d'objets saillants:	20
II.4.La Similitudes avec faculté du système visuel humain à être attiré par certains objet particulier d'une scène :	21
II.5.En quoi consiste la détection d'objets saillants ?	21
II.6.État de l'art sur la détection d'objets saillants	22
II.6.1. Méthodes basées sur l'estimation du contraste local:	22
II.6.2. Méthodes basées sur l'estimation du contraste global :	25
II.6.3.Méthodes basées sur la distance colorimétrique :	27
II.6.4.Méthodes d'estimation de la saillance dans le domaine fréquentiel:	28
II.7.Conclusion:	29
CHAPITRE III: Application ,Résultat et discussion :	30
III.4.1.Evaluation de la pertinence des caractéristiques :	31
III.4.2 Segmentation par diffusion graphique :	32
III.4.3.Evaluation des performances :	35
III.4.3.1Classification des pixels :	35
III.4.3.2.Mesures d'évaluation comparatives :	35
Tableau de comparaison entre les cartes de saillance et masque de quelques images:	37
III.5.Résultats expérimentaux:	38
III.5.1.Présentation des résultats:	38

III.6.Conclusion :.....	43
conclusion générale :.....	44

Table des figures

Figure 1 représentation de l'espace RVB.....	5
Figure 2:la représentation de l'espace CIE Lab	5
Figure 3 :La représentations cartésiennes et polaires de l'espace CIE Lab.	6
Figure 4:contour d'une image.....	8
Figure 5:La représentation de l'histogramme d'une image	8
Figure 6:Image binaire	9
Figure 7:Image Monochrome (256 couleurs).....	10
Figure 8 :Image polychrome (65536 couleurs) [4].	10
Figure 9:Principe du filtre média.	12
Figure 11:Principe du filtre minimum [5].	13
Figure 12: (a) Image originale en couleur (b) Image en niveaux de gris	14
Figure 14:Segmentation en régions par watershed	15
Figure 13:Extraction des contours d'une image à partir de ses gradients.....	15
a b.....	16
Figure 15:Segmentation en régions par k-means en 3 régions (a) et 4 régions (b).....	16
Figure 17: Une segmentation en superpixels d'une image d'entrée [16].....	17
Figure 16:Segmentation par contour actif. (a) Le contour initial et (b) le contour à l'état de convergence.....	17
figure 18:Quelque exemple d'objets d'intérêt.....	21
figure 19:Exemple d'un objet saillant	22
figure 20:exemple sur l'estimation du contraste local	23
Figure 21 :Structure pyramidale dont chaque niveau correspond à une échelle de représentation. La différence de gaussiennes est utilisée pour estimer le contraste centre/périphérie à travers la pyramide [58].....	24
Figure 22 :Limite des méthodes basées sur l'estimation du contraste local et celles basées sur l'estimation du contraste global.	29
Figure 23: quelques exemple d'images de la base de données	31
Figure 24 : Des exemples illustrant les avantages du traitement de mise à jour de la saillance [54]. (a) l'image en couleur, (b) carte de saillance générée sans le traitement de mise à jour, (c) carte de saillance générée avec le traitement de mise à jour.	34
Figure 25 : B:résultat de segmentation en superpixel de l'imageA	38
Figure 26:la carte de saillance obtenue de l'image C	39
Figure 31 :la carte saillance obtenue de l'image O.....	41
Figure 32 : R:résultat obtenue de la segmentation en superpixel de l'image Q	42
Figure 33 :T:la carte de saillance obtenue de l'image S.....	42

Remerciements :

tout d'abord merci ALLAH de nous avoir donné la patience et le courage d'accomplir ce travail. Nous tenons à exprimer notre profonde gratitude et nos sincères remerciements à notre encadreur : « Mr: Filali idir » Pour son aide précieuse et ses conseils et pour sa confiance au cours de l'élaboration de ce travail avec beaucoup de générosité durant cette période difficile d'épidémie . aussi à exprimer nos remerciements les plus respectueux à notre honorable jury.

Et enfin, que nos chers parents et familles, et bien avant tout, trouvent ici l'expression de nos remerciements les plus sincères et les plus profonds en reconnaissance de leurs sacrifices, aides, soutien et encouragement afin de nous assurer cette formation dans les meilleures conditions.

Nous remercions chaleureusement nos camarades de promotion pour les moments agréables que nous avons partagés ensemble.

Et enfin nous remercions tous ceux qui ont participé de près ou de loin à la réalisation de ce modeste mémoire.

Dédicace

*Au nom du dieu Clément et Miséricordieux louange à ALLAH
le tout puissant.*

*je dédie ce modeste travail en signe de respect,
reconnaissance et de remerciement :*

*A mes chères parents qui ont été toujours avec moi et grâce a
leur prières que je suis la aujourd'hui*

*A mes frères Kamel et Khaled ,et mes sœurs Nacera et Kahina
pour leurs encouragements et à toute ma famille*

A ma binôme Karima

a toutes mes chères amies Kahina ,Zahra ,Djouher,Malika ...

*A tous ceux qui m'ont encouragé et aidé tout au long de mes
études*

Amina

Dédicace

*Au nom du dieu Clément et Miséricordieux louange à ALLAH
le tout puissant.*

Je dédie ce modeste travail

*A mes très cher parents pour leur amour inestimable,
leur confiance, leur soutien, leurs sacrifices et toutes les
valeurs qu'ils
ont su m'inculquer.*

A mon cher mari Karim qui a été toujours à mes côtés

A mes frères: Kaci, Zohir, Nourdine et Morad ;

*A ma sœurs Hakima, Nassima, Fatiha, Djahida et Sara ;
qui m'ont apporté leurs encouragé et leurs soutiens .*

A ma binôme Amina ;

A toutes mes amies: Kahina, Djouher, Siham.....

*A tous ceux qui m'ont encouragé et aidé tout au long de mes
études*

Karima

Introduction générale :

La reconnaissance d'objets est une technique de vision par ordinateur permettant d'identifier des objets dans des images ou des vidéos. La détection d'objets est un résultat clé des algorithmes d'apprentissage en profondeur et d'apprentissage automatique.

Lorsque les humains regardent une photographie ou regardent une vidéo, ils peuvent facilement repérer des personnes, des objets, des scènes et des détails visuels. Le but est d'apprendre à un ordinateur à faire ce qui vient naturellement aux humains: acquérir un niveau de compréhension de ce que contient une image.

La reconnaissance d'objets est une technologie clé derrière les voitures sans conducteur, leur permettant de reconnaître un panneau d'arrêt ou de distinguer un piéton d'un lampadaire. Il est également utile dans une variété d'applications telles que l'identification de maladies dans la bioimagerie, l'inspection industrielle et la vision robotique.

La segmentation d'image est l'opération la plus importante dans un système de traitement d'images, car elle est située à l'articulation entre le traitement et l'analyse des images. L'intérêt de la segmentation est de partitionner une image en plusieurs régions homogènes, au sens d'un critère fixé a priori. L'intérêt de disposer de régions homogènes est de fournir des données simplifiées qui facilitent la tâche d'un système de reconnaissance de formes, ou autre système d'extraction des objets contenus dans l'image.

La détection des objets est parmi les problèmes clé en traitement d'image . La décision est très complexe dû à la diversité de leur format, taille et couleur et de la frontière mal défini entre les objets ainsi que la situation des images .

L'objectif du travail présenté dans ce mémoire consiste à implémenter par MATLAB un système pour détection des objets dans une image en utilisant la segmentation par diffusion graphique . Pour cela nous avons structuré notre mémoire autour de trois chapitres :

Chapitre I : Dans ce chapitre nous présentons un aperçu général sur le traitement d'image qui comporte l'acquisition d'image, ses caractéristique, ses types puis on a fait une introduction a la segmentation en général .

Chapitre II : ce chapitre est consacré a la détection des objets saillants et les différentes méthodes utilisés dans ce domaine .

Chapitre III : ce chapitre, est dédié à notre contribution d'implémentations de notre méthode sous MATLAB, résultats et discussions afin de faciliter la détection des objets saillants, avec un test d'évaluation du programme.

Enfin, on termine ce mémoire par une conclusion générale qui présente les principaux résultats et les perspectives de ce travail.

CHAPITRE I :

Traitement d'image et segmentation

I.1.Introduction:

Le traitement d'image peut être défini comme l'ensemble des méthodes et techniques opérant sur l'image afin d'extraire les informations les plus pertinentes ou tout simplement pour fournir une image plus perceptible à l'œil humain.

Le but de l'analyse d'image réside dans la description (forme, couleur, texture) et/ou la quantification (nombre, densité) des divers objets qui composent l'image étudiée. L'analyse trouve sa place dans de nombreux domaines tels que la robotique, les sciences de la vie ou encore la sécurité[1].

La segmentation est une des étapes critiques de traitement qui contient plusieurs approches Dans ce chapitre nous présentons quelques notions de base du domaine de traitement d'image numérique tels que : la définition d'image, caractéristiques d'image : *Pixels* ,espaces de couleur, texture, luminance, contours et les types d'image ,histogrammes et opération de filtrages, et pour la segmentation on cite l'approche contour, l'approche région, segmentation en superpixels et segmentation basé sur les bassins versant (watershed segmentation).

I.2Traitement d'une image

I.2.1Définition d'une image :

L'image est une représentation d'une personne ou un animal ou des choses se trouvant dans l'environnement. C'est aussi un ensemble structuré d'informations qui, après affichage sur l'écran, ont une signification pour l'œil humain.

Une image numérique est une matrice de pixels repérés par leur coordonnées (x, y).S'il s'agit d'une image couleur, un pixel est codé par 3 composantes (r, g, b), représentant respectivement les doses de rouge, vert et bleu qui caractérisent la couleur du pixel.

S'il s'agit d'une image en niveau de gris, il est codé par 1 composante comprise au sens large entre 0 et 255, représentant la luminosité du pixel.

I.2.2Caractéristique d'image :

L'image est caractérisé par les paramètres suivants :

I.2.2.1Pixels:

Une image est constituée d'un ensemble de points appelés pixels (pixel est une abréviation de *PIC*ture *E*lement) . Le pixel représente ainsi le plus petit élément constitutif d'une image numérique, L'ensemble de ces pixels est contenu dans un tableau à deux dimensions constituant l'image, qui renvoie toute l'information qui constitue l'image dans son intégralité[3].

I.2.2.2La dimension:

La dimension d'une image correspond à sa taille en terme du nombre de lignes et de colonnes la constituant.

1.2.2.3la résolution :

La résolution, terme souvent confondu avec la "définition", détermine par contre le nombre de points par unité de surface, exprimé en *points par pouce (PPP*, en anglais **DPI** pour *Dots Per Inch*); un pouce représentant 2.54 cm. La résolution permet ainsi d'établir le rapport entre le nombre de pixels d'une image et la taille réelle de sa représentation sur un support physique. Une résolution de 300 dpi signifie donc 300 colonnes et 300 rangées de pixels sur un pouce carré ce qui donne donc 90000 pixels sur un pouce carré[4].

1.2.2.4le codage de la couleur :

Pour représenter informatiquement une image, il suffit de créer un tableau de pixels dont chaque case contient une valeur. La valeur stockée dans une case est codée sur un certain nombre de bits déterminant la couleur ou l'intensité du pixel, on l'appelle profondeur de codage (parfois *profondeur de couleur*). Il existe plusieurs standards de codage de la profondeur : bitmap noir et blanc, bitmap 16 couleurs ou 16 niveaux de gris, bitmap 256 couleurs ou 256 niveaux de gris, palette de couleurs (colormap) et Couleurs vraies (truecolor)) ou « *couleurs réelles* » : cette représentation permet de représenter une image en définissant chacune des composantes (RGB, pour rouge, vert et bleu). Chaque pixel est représenté par un entier comportant les trois composantes, chacune codée sur un octet, c'est-à-dire au total 24 bits (16 millions de couleurs). Il est possible de rajouter une quatrième composante permettant d'ajouter une information de transparence ou de texture, chaque pixel est alors codé sur 32 bits[4].

1.2.2.5luminance :

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface. Pour un observateur lointain, le mot luminance est substitué au mot brillance, qui correspond à l'éclat d'un objet [5].

1.2.2.6Contraste :

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'image.

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 :

$$C = \frac{L1-L2}{L1+L2} [5]..... I .1$$

1.2.2.7Format :

Un format d'image est une représentation informatique de l'image, associée à des Informations sur la façon dont l'image est codée et fournissant éventuellement des indications sur la manière de la décoder et de la manipuler [6].

il existe un grand nombre de formats d'image et on site ici les plus utilisés:

format	Compression des données	Nb de couleurs	Affichage progressif	Usage
--------	-------------------------	----------------	----------------------	-------

BMP	Non compressé	De 2 a 16 millions	Non	Image non dégradée mais très lourde. Stockage
JPEG	Réglable, avec prête de qualité. Plus la compression est importante ,plus l'image est dégradée	16 millions	Oui	Tous usages ,selon compression. Images « naturelles ».
GIF	Oui, sans perte de qualité	De 2 à 256 avec palette.	Oui	Logos et Internet. Supporte les animations et la transparence
TIFF	Réglable ,au choix sans perte ou avec perte de qualité	16 millions	Non	Tous sauf Internet
PNG	Oui ,sans perte de qualité	De 2 à 256 ou 16 millions	Oui	Tous , recommandé Internet mais incompatible avec les navigateur anciens. Supporte la transparence

Tableau 1 :les principaux formats d'une image .

1.2.2.8Espace de couleurs :

1.2.2.8.1Représentation des couleurs:

Une couleur est généralement représentée par trois composantes. Ces composantes définissent un espace des couleurs. On peut citer l'espace RVB, l'espace CIE XYZ ou Yxy, ou encore l'espace Lab. Selon l'espace de couleurs choisi pour représenter une image couleur, le nuage des couleurs (c'est à dire l'ensemble des couleurs de l'image) n'aura pas la même répartition dans l'espace 3D. Les espaces de couleurs classiques, tels que le RVB, CIE XYZ, etc ..., sont issus d'une approche purement physique, sans prise en compte de données psychophysiques. Dans le cas d'autre espaces de couleur, tels que l'espace Lab, l'approche physique est corrigée selon des données de la vision humaine .

1.2.2.8.2L'espace RVB:

L'espace RVB est sans doute l'espace de couleurs le plus utilisé. Les systèmes de télévision s'y appuient fortement parce qu'ils renvoient la couleur réelle d'un objet tel que l'homme la perçoit . La représentation des couleurs dans cet espace donne un cube appelé cube de Maxwell [7].

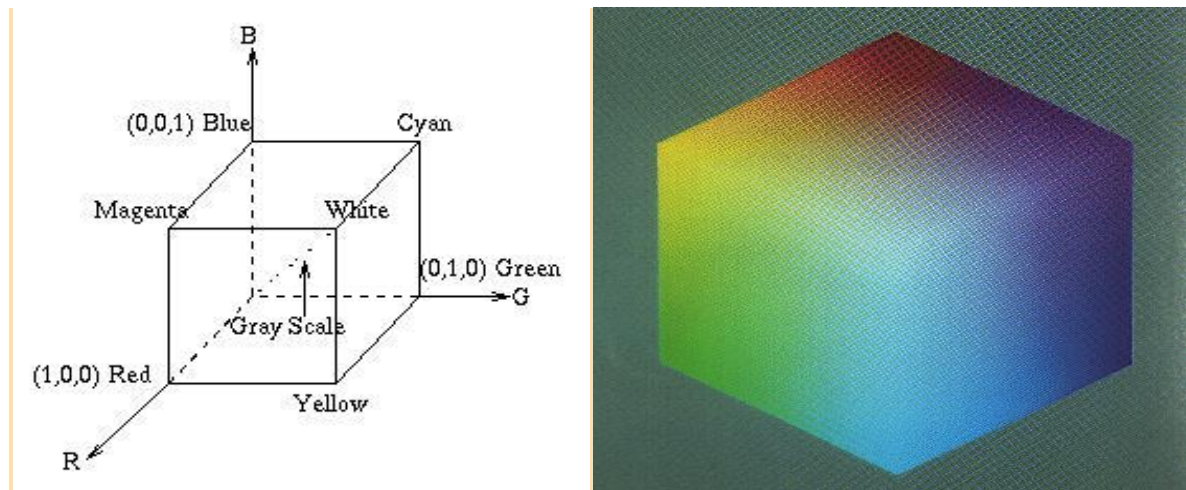


Figure 1 représentation de l'espace RVB

I.2.2.8.3L'espace CIE Lab :

Le système Lab est issu du CIE XYZ. Il essaye de prendre en compte la réponse logarithmique de l'œil, il possède le grand avantage d'être uniforme. Il est très utile dans le cas de mélanges de pigments, par exemple, pour l'industrie graphique ou du textile.

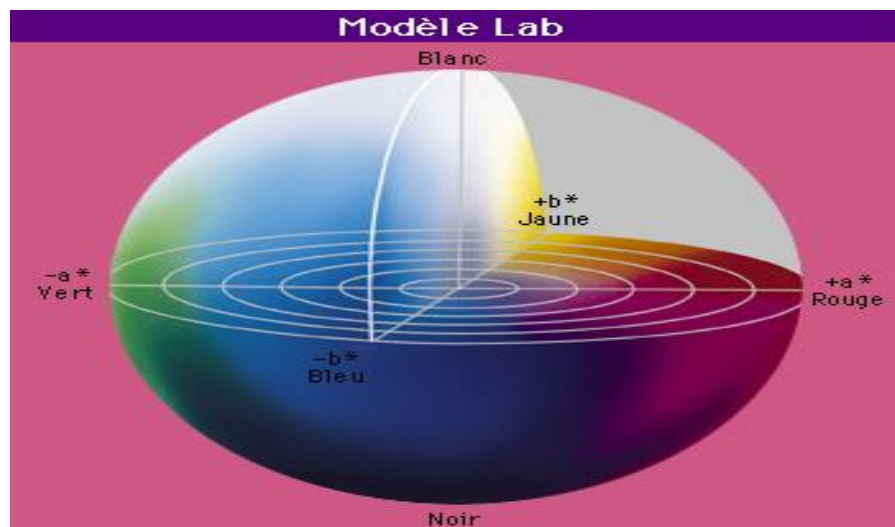


Figure 2: la représentation de l'espace CIE Lab

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

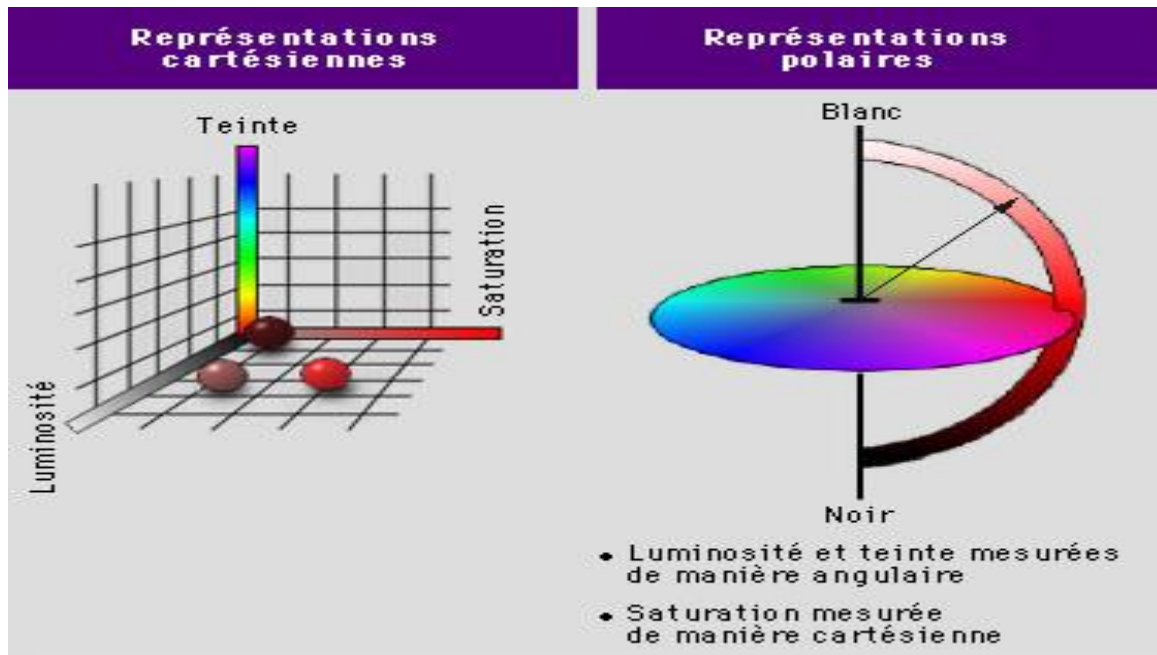


Figure 3 :La représentations cartésiennes et polaires de l'espace CIE Lab.

I.2.2.8.4 Opérations de passage des espaces RVB et Lab:

L'espace CIE XYZ est une étape intermédiaire incontournable de la conversion. L'espace Lab est en effet défini par rapport à l'espace XYZ.

a. De l'espace RVB à l'espace Lab

La première étape consiste à passer des composantes RVB aux composantes XYZ. On utilise pour cela une matrice de conversion

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.618 & 0.177 & 0.205 \\ 0.299 & 0.587 & 0.114 \\ 0 & 0.056 & 0.944 \end{pmatrix} * \begin{pmatrix} R \\ V \\ B \end{pmatrix} \dots \dots \dots I.2$$

Ensuite, il s'agit de passer de l'espace XYZ à l'espace Lab. Pour cela, on utilise les formules de conversion suivantes :

$$L = 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \quad \text{pour} \quad \frac{Y}{Y_n} > 0.008856 \dots \dots \dots I.3$$

$$L = 903.3 \frac{Y}{Y_n} \text{pour} \quad \frac{Y}{Y_n} \leq 0.008856 \dots \dots \dots I.4$$

$$a = 500 \left(f \left(\frac{X}{X_n} \right) - f \left(\frac{Y}{Y_n} \right) \right) \dots \dots \dots I.5$$

$$b = 200 \left(f \left(\frac{Y}{Y_n} \right) - f \left(\frac{Z}{Z_n} \right) \right) \dots \dots \dots I.6$$

où

pour $t > 0.008856$ $f(t) = \sqrt[3]{t} \dots \dots \dots I.7$

pour $t \leq 0.008856$ $f(t) = 7.7787t + \frac{16}{116} \dots \dots \dots I.8$

Xn, Yn et Zn correspondent au blanc décrit dans l'espace XYZ. On les obtient pour RGB=(255,255,255).

b. De l'espace Lab à l'espace RVB

On convertit les composantes Lab dans l'espace XYZ. Pour cela, on inverse les formules précédentes:

$Y = Yn \cdot \frac{L}{903.3} \dots \dots \dots \text{pour } L \leq 8 \dots \dots \dots I.9$

$Y = Yn \cdot \left(\frac{L + 16}{116}\right)^3 \dots \dots \dots \text{pour } L > 8 \dots \dots \dots I.10$

$X = Xn \cdot f\left(\frac{a}{500} + \frac{L + 16}{116}\right) \dots \dots \dots I.11$

$Z = Zn \cdot f\left(\frac{L}{116} - \frac{b + 16}{200}\right) \dots \dots \dots I.12$

où

pour $t > 0.207$ $f(t) = t^3 \dots \dots \dots I.13$

pour $t \leq 0.207$ $f(t) = \frac{116t - 16}{903.3} \dots \dots \dots I.14$

Xn, Yn, Zn sont les composantes du blanc.

Ensuite, par application de la matrice inverse (RVB->XYZ), on obtient les composantes RVB:

$$\begin{pmatrix} R \\ V \\ B \end{pmatrix} = \begin{pmatrix} 1.876 & -0.533 & -0.343 \\ -0.967 & 1.998 & -0.031 \\ 0.057 & -0.118 & 1.061 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \dots \dots \dots I.15$$

I.2.2.8.5Ecart de couleur

Lorsqu'il s'agit de travailler sur les couleurs d'une image, comme dans le cas d'une quantification, deux couleurs qui sont proches dans l'espace de couleur (au sens de la distance euclidienne le plus souvent), peuvent paraître assez différentes pour l'oeil, ce qui est le cas pour l'espace des couleurs RGB.

Par contre, dans l'espace LAB qui est uniforme, deux couleurs proches en distance le sont

aussi pour l'œil [7].

1.2.2.9 Contours et textures

Les contours représentent la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative. Les textures décrivent la structure de ceux-ci. L'extraction de contour consiste à identifier dans l'image les points qui séparent deux textures différentes



Figure 4: contour d'une image

1.3.2.10 Histogramme

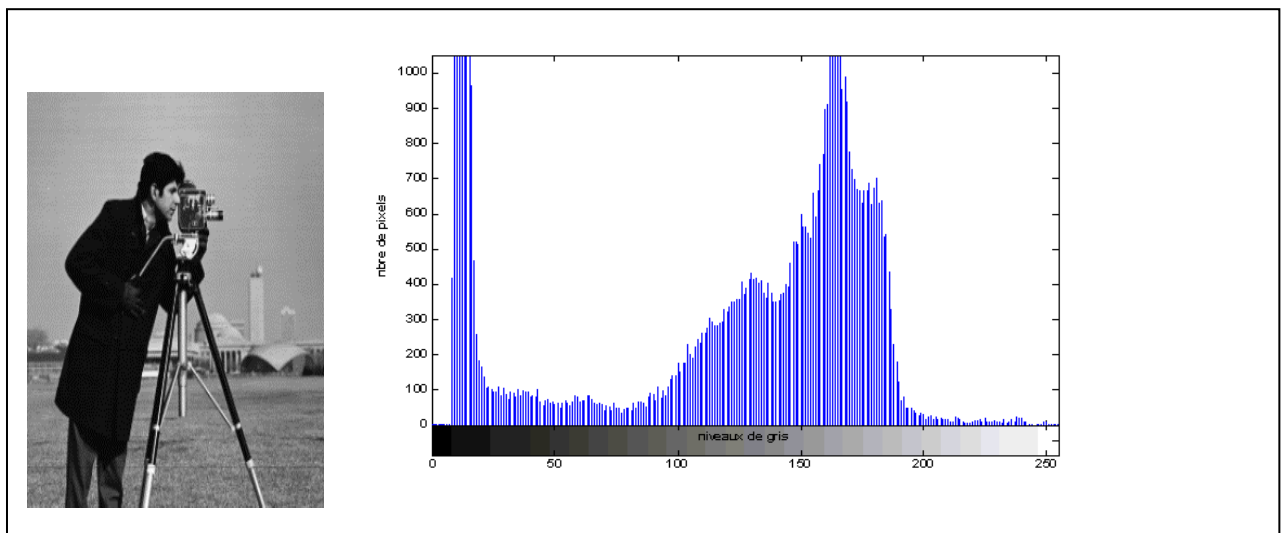


Figure 5: La représentation de l'histogramme d'une image

L'histogramme d'une image est une fonction discrète. Elle représente le nombre de pixels en fonction du niveau de gris.

Lorsque cette fonction est normalisée entre 0 et 1 pour tous les niveaux de gris, on peut la voir comme une densité de probabilité qui fournit la probabilité de trouver un certain niveau de gris de l'image. Ainsi le niveau de gris d'un pixel devient une variable aléatoire dont la valeur

dépend du résultat d'une expérience aléatoire sous-jacente. D'où un traitement statistique des images.

I.3.3 Les types des images :

On distingue trois types d'images :

- **Binaire** : 2 couleurs (arrière-plan et avant-plan).
- **Monochrome** : variations d'une même teinte.
- **Polychrome** : " vraies " couleurs.

I.3.3 .1Images binaires (en noir et blanc) :

Les images binaires sont les plus simples. Bichromes (dont les pixels ne peuvent avoir que les valeurs 0 et 1). Le 0 correspond à un pixel noir et 1 à un pixel blanc. Le niveau de gris est donc codé sur un seul bit.

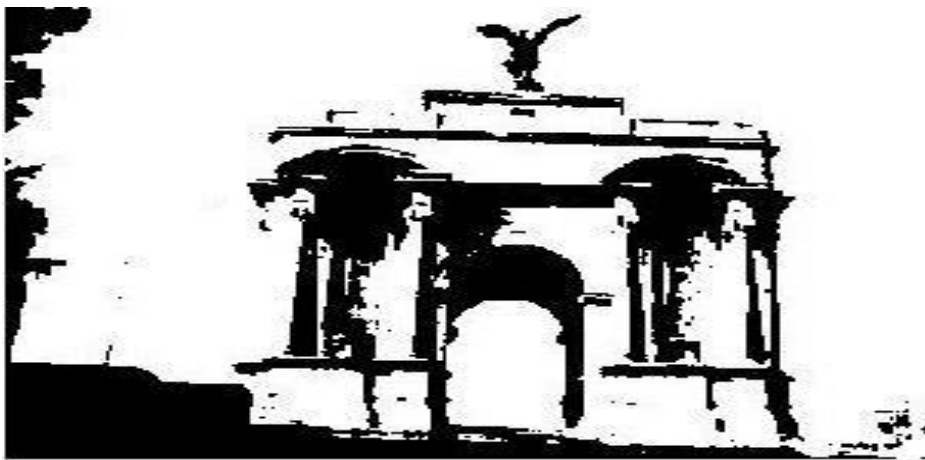


Figure 6:Image binaire

I.3.3 .2Images à niveaux de gris (Monochromes) :

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires. Donc pour représenter les images à niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise par exemple entre 0 et 255. Chaque pixel n'est donc plus représenté par 1 bit, mais par 1 octet. Pour cela, il faut que le matériel utilisé pour afficher l'image, soit capable de produire les différents niveaux de gris correspondant.

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



Figure 7:Image Monochrome (256 couleurs).

I.3.3 Images en couleurs (Polychromes) :

Elle est obtenue par la combinaison de trois couleurs dites primaires : rouge, vert et bleu(RVB). Chaque couleur est codée comme une image à niveaux de gris, avec des valeurs allant de 0 à 255. Ainsi, pour $R=V=B=0$ nous aurons un noir pur, et pour $R=V=B=255$ nous aurons un blanc pur. La représentation des images couleurs se fait donc soit par une image dont la valeur du pixel est une combinaison linéaire des valeurs des trois composantes couleurs, soit par trois images distinctes représentant chacune une composante couleur.

On distingue généralement deux types d'images:

- **Images** 24 bits ou « **couleurs vraies** »
- **Images** à palettes, **images** en 256 **couleurs** (8 bits)



Figure 8 :Image polychrome (65536 couleurs) [4].

I.3.4 Opération de filtrages

Les images numériques telles qu'elles sont acquises, sont très souvent inexploitable pour le traitement d'images. Elles contiennent des signaux bruités. Pour remédier à cela, différents prétraitements pour l'amélioration ou la correction sont effectués.

On peut scinder les filtres en deux grandes catégories :

I.3.4 .1Filtres linéaires

Les filtres linéaires transforment un ensemble de données d'entrée en un ensemble de données de sortie par une convolution bidimensionnelle qui est une opération mathématique. Ils permettent de supprimer le bruit dans l'image. Chaque filtre a une taille $N \times N$ avec N est un nombre impair.

Les filtres linéaires les plus connus sont les filtres passe-bas et passe-haut.

I.3.4 .1.1 Filtre passe-bas (lissage) :

Ce filtre n'affecte pas les composantes de basse fréquence dans les données d'une image, mais doit atténuer les composantes de haute fréquence. L'opération de lissage est souvent utilisée pour atténuer le bruit et les irrégularités de l'image. Elle peut être répétée plusieurs fois, ce qui crée un effet de flou. En pratique, il faut choisir un compromis entre l'atténuation du bruit et la conservation des détails et contours significatifs.

$1/9^*$

1	1	1
1	1	1
1	1	1

Masque du filtre

I.3.4 .1.2 Filtre Passe-haut (Accentuation) : Le renforcement des contours et leur extraction s'obtiennent dans le domaine fréquentiel par l'application d'un filtre passe-haut. Le filtre digital passe-haut a les caractéristiques inverses du filtre passe-bas. Ce filtre n'affecte pas les composantes de haute fréquence d'un signal, mais doit atténuer les composantes de basse fréquence

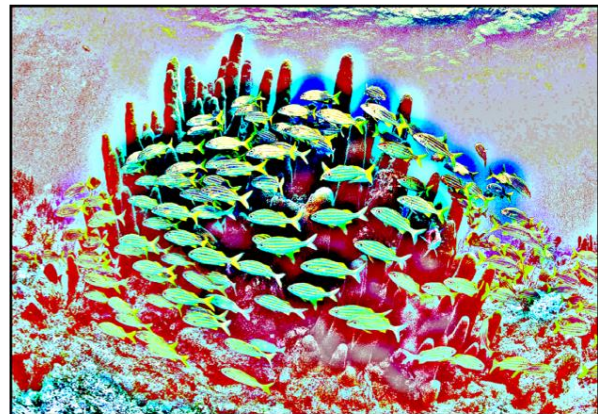
H=

-1	-1	-1
-1	9	-1
-1	-1	-1

Masque de filtre



Image d'origine



Filtre Passe-haut appliqué

I.3.4 .2 Filtres non linéaire

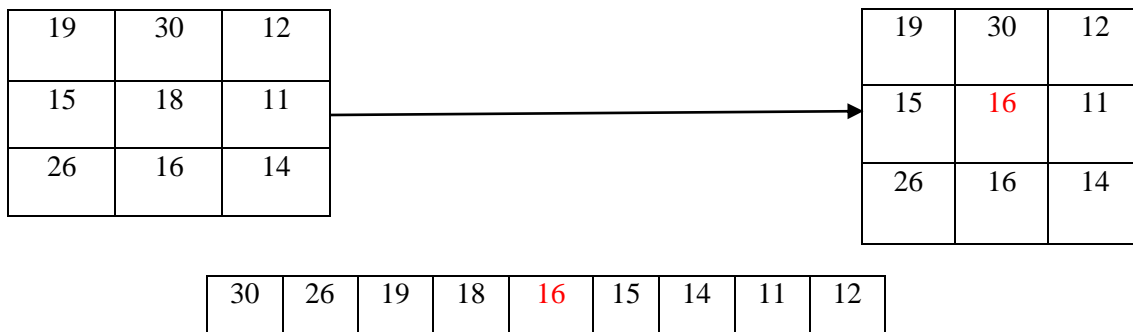
Ils sont conçus pour régler les problèmes des filtres linéaires, Leur principe est le même que celui des filtres linéaires, il s'agit toujours de remplacer la valeur de chaque pixel par la valeur d'une fonction calculée dans son voisinage. La différence majeure, est que cette fonction n'est plus linéaire mais une fonction quelconque (elle peut inclure des opérateurs de comparaisons ou de classification). Les filtres non linéaires les plus connus sont :

I.3.4 .2 .1 Filtre médian

Ce filtre est très utilisé pour éliminer le bruit sur une image qui peut être de différentes origines (poussières, petits nuages, baisse momentanée de l'intensité électrique sur les capteurs, ...). L'avantage de ce filtre réside dans le fait qu'il conserve les contours alors que les autres types de filtres ont tendance à les adoucir.

L'algorithme de filtre médian est le suivant :

1. Trier les valeurs par ordre croissant.
2. Remplacer la valeur du pixel centrale par la valeur située au milieu de la triée.
3. Répéter cette opération pour tous les pixels de l'image [8].



Original Image



with Median Filter

Figure 9: Principe du filtre médian.

I.3.4 .2 .2 Filtre maximum

On applique le même traitement que celui du filtre médian mais la valeur du pixel du centre comme on le montre ici va être changée par le maximum.

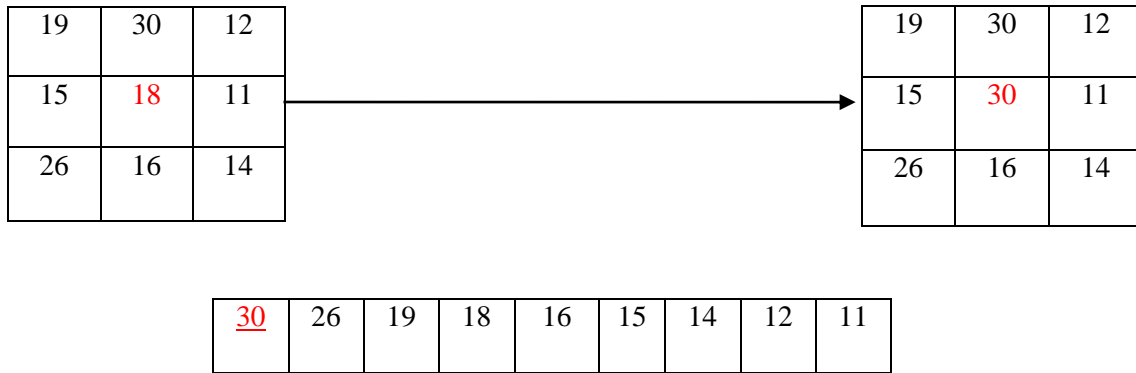


Figure 10: Principe du filtre maximum

I.3.4 .2 .3 Filtre minimum

On applique le même traitement que celui du filtre maximum mais, cette fois, la valeur du pixel du centre comme la montre la figure va être remplacée par le minimum.

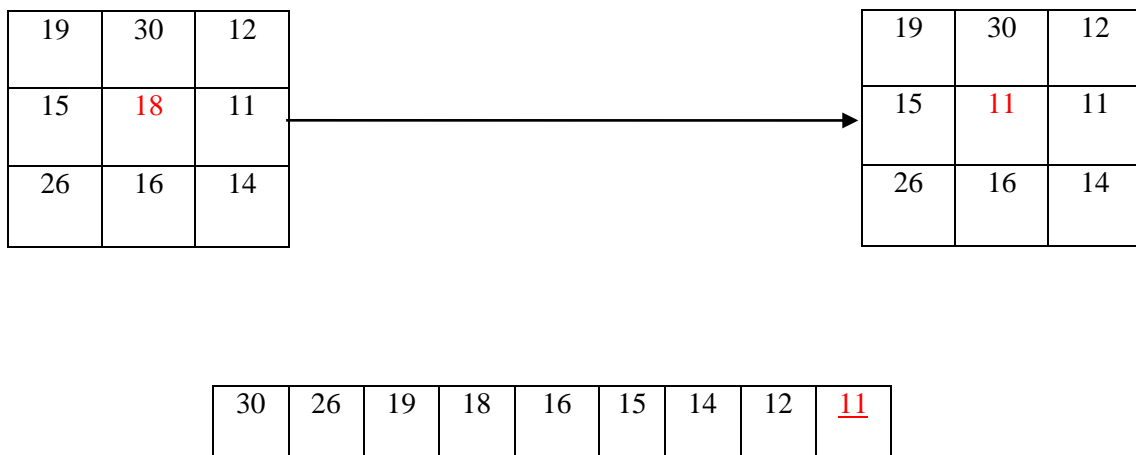


Figure 11: Principe du filtre minimum [5].

II. Segmentation d'images :

II.1 Définition :

Les méthodes de segmentation s'appliquent sur des images matricielles, afin d'extraire une ou plusieurs zones d'intérêt connexes, qui ont un pouvoir discriminant spécifique vis-à-vis du domaine applicatif.

Les techniques de segmentation diffèrent selon le type d'application et selon l'information recherchée dans l'image. Les zones extraites peuvent être des points et leur voisinage, des lignes ou des régions. Prenons par exemple l'image de la figure 12 et essayons, dans les sections qui suivent, de l'analyser et de segmenter le papillon. La figure 12 (a) représente l'image originale en couleur et la figure 12(b) représente la même image en niveaux de gris.

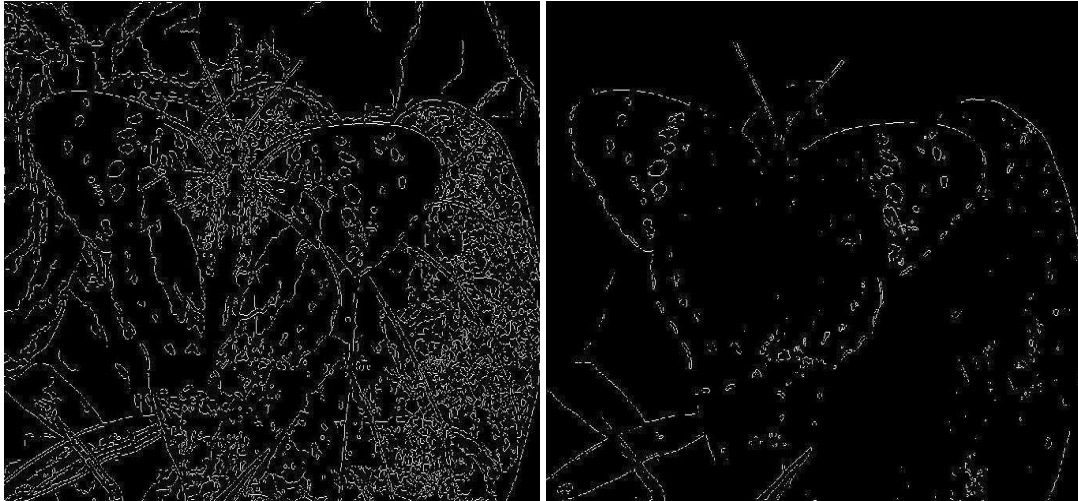


Figure 12: (a) Image originale en couleur (b) Image en niveaux de gris

La segmentation d'images peut être classifiée en général a des approches qui cherchent à localiser les contours dans l'image, d'autres à extraire les régions dans l'images ou bien des approches qui détecte plutôt les objets dans l'image ou une forme donnée.

II.2.Approche contour:

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. Le principe de la détection de contours repose donc sur l'étude des dérivées de la fonction d'intensité dans l'image. Citons par exemple les approches basées sur les différences finies comme les opérateurs du gradient et du laplacien, les filtres de Prewitt, Sobel, Roberts ; ou basées sur des critères d'optimalité comme le filtre de Canny [9]. Signalons qu'une présence du bruit dans l'image peut rendre la tâche de la détection des contours un peu difficile. Les méthodes exploitant des accumulateurs, telles que la transformée de Hough, ont été proposées dans d'autres travaux pour extraire des courbes paramétrées (droites, cercles, ellipses) de manière robuste ; mais ces méthodes sont généralement assez lourdes, surtout si le nombre de paramètres augmentent (deux pour une droite, quatre pour une ellipse, etc.) .



(a) Filtre de Canny

(b) Filtre de Prewitt

Figure 13:Extraction des contours d'une image à partir de ses gradients..

La Figure 13 illustre deux extractions possibles des gradients en utilisant les filtres de Canny (figure 13(a)) ou le filtre de Prewitt (figure 13(b)). Ces gradients définissent des contours qui n'ont pas tous le même intérêt pour nous. Les contours de la feuille sont bien marqués, alors que le contour du papillon n'a pas une intensité constante. Si l'on peut déterminer l'intensité à laquelle les gradients sont détectés, on ne peut cependant pas déterminer leur nombre, s'assurer de leur connexité, ou leur donner un sens quant à leur appartenance à l'objet. Autrement dit, tous les gradients détectés n'appartiennent pas au contour du papillon, et nous ne saurions les faire disparaître sans perdre de la même manière une partie des contours du papillon. De plus, ces contours ne sont pas forcément fermés et ne définissent donc pas des régions séparées.

II.3.Segmentation basée sur région :

Certaines méthodes morphologiques, comme la segmentation par la ligne de partage des eaux [10], permettent de définir des régions à partir des gradients de l'image. La figure 14 illustre la segmentation de l'image de la figure 13 en utilisant la méthode de la ligne de partage des eaux (*watershed*). Cette méthode est rapide en temps de calcul mais fournit souvent un nombre très grand de régions qu'il faudra par la suite fusionner pour obtenir une segmentation correcte des objets dans l'image.

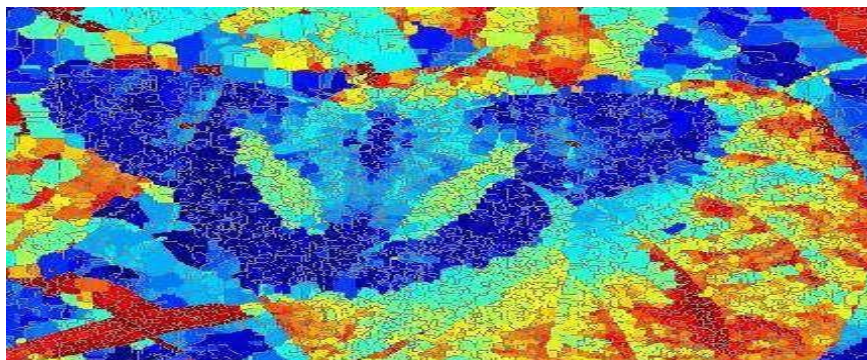


Figure 14:Segmentation en régions par watershed

Une autre approche est appelée division-fusion (*split and merge*) [11] et consiste soit à diviser une image et puis fusionner les régions adjacentes selon des critères d'homogénéité, soit à diviser l'image de façon itérative tant que les régions ne sont pas assez homogènes. On utilise souvent des arbres ou des graphes (*quadtree*) pour effectuer cette étape.

Une autre méthode basée région dite croissance de régions (*region-growing*) [12], est classifiée parmi les méthodes de segmentation basée sur les pixels vu qu'elle fait agrandir un ensemble de semences initiales (des pixels initiaux) choisies dans l'image. Cela rend cette méthode assez sensible au nombre et au positionnement des semences initiales. Cette approche examine (selon un critère d'homogénéité) le voisinage de ces pixels initiaux afin de déterminer si le voisinage du pixel peut être inclus dans la région contenant ce pixel et ainsi faire croître cette région.

Il existe des méthodes qui permettent une classification (*clustering*) des pixels de l'image en sous-ensembles possédants des caractéristiques proches. Au lieu de sélectionner un seul seuil pour répartir les pixels en deux classes, les histogrammes (pour une image couleur) sont analysés pour extraire plusieurs sous-classes [13]. L'approche présentée dans [13] a été en particulier proposée au LAAS par G. AviñaCervantés comme étape préliminaire de la segmentation d'images couleur.

Une autre approche très connue est la méthode de segmentation par *k-means* ([14], [15]), qui consiste à classifier les pixels de l'image en k classes en minimisant, pour chaque classe, une distance entre un pixel de la classe et le représentant de cette classe.

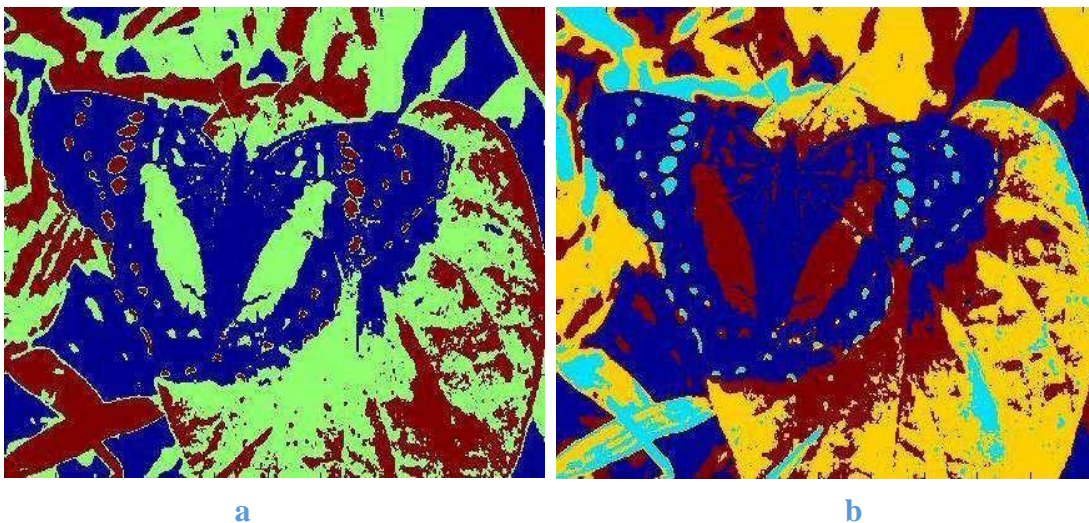


Figure 15: Segmentation en régions par k-means en 3 régions (a) et 4 régions (b)

La figure 15 présente la segmentation de l'image du papillon en 3 régions (figure 15(a)) et en 4 régions (figure 15 (b)). Pourtant, cette méthode ne permet pas de distinguer les pixels faisant partie de l'objet (qu'on veut segmenter) de ceux appartenant au fond de l'image. Le papillon dans la figure 15 est constitué de plusieurs régions de différentes couleurs et il n'est pas distingué du reste de l'image.

II.4. Segmentation par contour actif

Une dernière approche que nous présentons consiste à définir un contour fermé et à faire évoluer vers l'objet d'intérêt. Les déformations de cette courbe sont déduites de la

dérivation d'une fonctionnelle à optimiser caractérisant l'objet. Les travaux sur les contours actifs se divisent en deux variantes qui sont des méthodes uniquement basées sur des informations contour, ou bien celles intégrant des caractéristiques sur les régions et donc des informations plus globales sur les objets à segmenter d'une image.

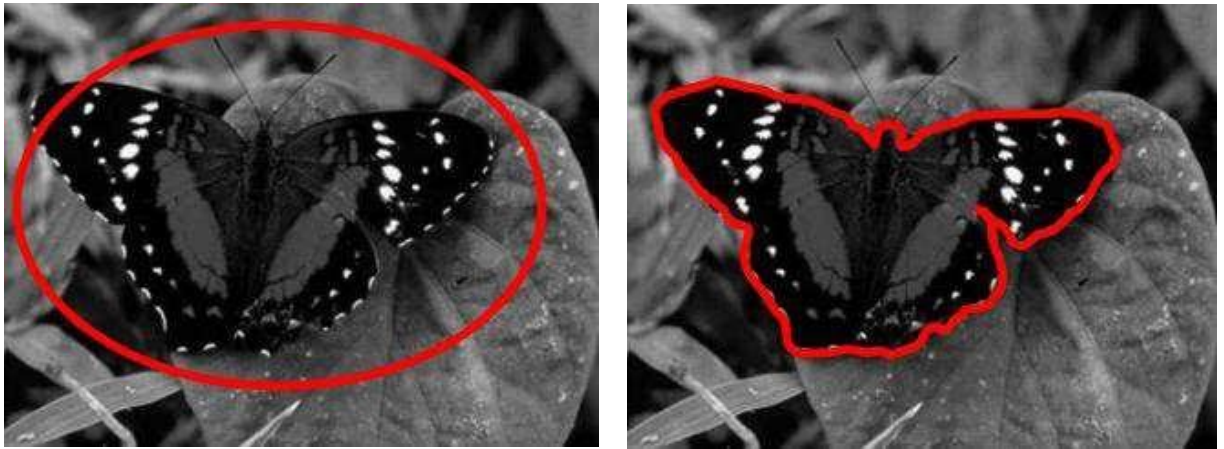


Figure 16: Segmentation par contour actif. (a) Le contour initial et (b) le contour à l'état de convergence

La figure 16 illustre la segmentation du papillon par contour actif. La position initiale et celle à l'état de convergence du contour actif sont illustrées dans les figure 16 (a) et 16 (b), respectivement.

II.5. Segmentation Superpixel

Les pixels d'image sont l'unité de base dans la plupart des tâches de traitement d'image. Cependant, ils sont la conséquence de la représentation discrète d'images et non d'entités naturelles. Les superpixels sont le résultat d'un regroupement perceptuel de pixels, ou inversement, les résultats d'une sur-fragmentation d'image. Les superpixels contiennent plus d'informations que les pixels et s'alignent mieux avec les bords de l'image que les correctifs d'image rectangulaires. Le superpixel peut entraîner une accélération substantielle du traitement ultérieur.

Une segmentation en superpixels d'une image d'entrée est illustrée dans la figure suivante:



Figure 17: Une segmentation en superpixels d'une image d'entrée [16]

Conclusion

Dans ce chapitre ,en premier lieu a présenté les notions de base du traitement d'image qui est l'ensemble des méthodes et techniques opérant sur celles-ci, dans le but de rendre cette opération possible, plus simple, plus efficace et plus agréable, d'améliorer l'aspect visuel de l'image et d'en extraire des informations jugées pertinentes. La définition qui y est développées est celles des caractéristiques et types d'image pour L'initiation aux traitements approfondis des images de prétraitement qui consiste à améliorer la qualité de nos image puis on a attardé sur la présentation de quelques méthodes de segmentation d'image les plus communes.

Chapitre II :

Détection d'objets saillants

II.1.Introduction

Les êtres humains possèdent la faculté d'identifier rapidement des objets saillants dans des scènes visuelles complexes avec une précision remarquable. La saillance d'une région dans une image est généralement caractérisée par sa rareté ou bien par le contraste local de ses attributs comme la couleur, la texture et les contours par rapport à son entourage [17]. Les premiers modèles de saillance ont été développés principalement à base de la prédiction de fixations oculaires dans les images naturelles où le but est de comprendre l'attention visuelle humaine [18,19,20]. Récemment, plusieurs méthodes ont été proposées pour détecter les objets saillants qui se distinguent de leur environnement [70,50,62]. Les approches de détection de saillance définissent l'unicité (rareté) d'une région dans un contexte spatiale local ou global. L'irrégularité locale se mesure dans un contexte spatial réduit à une région et son entourage. Par contre, l'irrégularité globale se mesure dans un contexte spatial plus large englobant toutes les régions de la scène visuelle. Nous allons dans ce chapitre nous baser sur ces contraintes spatiales pour classer les méthodes de détection d'objets saillants en trois grandes classes majeures: 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.

II.2.Qu'est-ce que la saillance?

Dans notre champ de vision, il y a toujours des choses qui se démarquent plus que d'autres. Si vous vous retrouvez à contempler une ville d'une hauteur par exemple, vous pouvez être attiré par un gratte-ciel à proximité, une lumière clignotante ou même un manteau rouge que quelqu'un porte en dessous.

La saillance est l'aspect de tout stimulus qui le distingue de la foule. La raison pour laquelle un stimulus particulier a une telle saillance peut être due au contraste, c'est-à-dire une ligne blanche sur un fond noir ou à la suite de facteurs émotionnels ou cognitifs.

Par exemple, nous pouvons nous concentrer sur quelque chose parce que nous le recherchons activement ou parce qu'il déclenche quelque chose dans notre passé ou notre mémoire.

La saillance est le plus souvent discutée en relation avec le système visuel, mais elle est utilisée par tous les systèmes perceptifs tels que le son et le toucher. Si nous avons faim, l'odeur d'un aliment préféré peut être très saillante par exemple.

Les mécanismes par lesquels les humains accordent à certains stimuli une concentration plus attentionnelle que d'autres tiennent probablement à notre passé évolutif. Nos ressources cognitives limitées nécessitent un moyen d'identifier les stimuli les plus pertinents pour l'apprentissage et / ou la survie.

Le monde est plein de stimuli partout où vous vous tournez et nous ne pouvons pas nous occuper de tout cela à la fois. Comment notre système visuel sait-il où se concentrer?

Afin d'être attentif au danger, le système sensoriel est attiré par des choses inhabituelles ou en conflit avec l'environnement. Si vous sursautez à un bruit fort pendant la nuit, votre système

sensoriel accorde ce bruit saillant. Cela revêt évidemment une importance historique pour la survie.

Bruit de saillance: La saillance peut également être apprise. Ceux qui parlent couramment une langue particulière, par exemple, peuvent être plus à l'écoute de certains sons spécifiques à cette langue.

Lorsque l'attention est dirigée par la saillance, elle est souvent de nature réflexive. Cela signifie que cela se produit instinctivement et sans beaucoup de conseils cognitifs.

Comme la façon dont votre main pourrait immédiatement se branler après avoir touché un poêle chaud, la saillance informe automatiquement où sont placées vos capacités d'attention et de perception.

Vous décidez rarement où chercher, d'autant plus que votre système visuel décide pour vous![23].

II.3 .Historique de la détection d'objets saillants:

L'un des premiers modèles de saillance, proposé par Itti et al. [24], a généré la première vague d'intérêt dans plusieurs disciplines, notamment la psychologie cognitive, les neurosciences et la vision par ordinateur. Ce modèle est une implémentation de cadres informatiques généraux et de théories psychologiques de l'attention ascendante antérieurs basés sur des mécanismes centraux (par exemple, la théorie de l'intégration des fonctionnalités de Treisman et Gelade [25], le modèle de recherche guidée de Wolfe et al. [26] , et l'architecture d'attention computationnelle de Koch et Ullman [27]). En réf. [24], Itti et al. Montrez quelques exemples où leur modèle est capable de détecter des discontinuités spatiales dans les scènes. Des enquêtes comportementales (par exemple, réf. [28]) et informatiques (par exemple, réf. [29]) ultérieures ont utilisé des fixations comme moyen de vérifier l'hypothèse de saillance et de comparer les modèles.

Une deuxième vague d'intérêt a surgi avec les travaux de Liu et al. [30, 31] and Achanta et al. [32] a détecté la saillance comme un problème de segmentation binaire. Ces auteurs se sont inspirés de certains modèles antérieurs s'efforçant de détecter des régions saillantes ou des proto-objets (par exemple, Ma et Zhang [33], Liu et Gleicher [34], et Walther and Koch [35]). Une multitude de modèles de science a alors émergé. Cependant, la façon dont cette nouvelle définition est liée à d'autres domaines de vision par ordinateur tels que la segmentation d'image (par exemple, réf. [36, 37]), la génération de propositions d'objet indépendantes de la catégorie (par exemple, réf. [38, 39, 40]), la fixation, est moins claire. la prédiction (par exemple, Réf. [29, 41–45]) et la détection d'objet (par exemple, Réf. [46, 47]). Une troisième vague d'intérêt est apparue récemment avec l'essor de la popularité des réseaux de neurones convolutionnels (CNN) [48], et en particulier avec l'introduction de réseaux de neurones entièrement convolutionnels [49]. éliminer le besoin de fonctionnalités fabriquées à la main, et réduire la dépendance à l'égard des connaissances sur les biais centraux, et ont donc été adoptées par de nombreux chercheurs. Un modèle basé sur CNN contient normalement des centaines de milliers de paramètres et de neurones accordables avec des tailles de champ de réception variables. Les neurones avec de grands champs récepteurs fournissent des informations globales qui peuvent aider à mieux identifier la région la plus saillante d'une image, tandis que les neurones avec de petits champs récepteurs fournissent des informations locales qui peuvent être utilisées pour affiner les cartes de saillance produites par les couches supérieures. Cela permet de mettre en évidence les régions saillantes et d'affiner leurs limites. Ces propriétés souhaitables permettent aux modèles basés sur CNN d'atteindre des

performances sans précédent par rapport aux modèles basés sur des fonctionnalités fabriqués à la main. Les modèles CNN deviennent progressivement la direction dominante de la détection d'objets saillants.

II.4. La Similitudes avec faculté du système visuel humain à être attiré par certains objet particulier d'une scène :

Notre environnement regorge d'objets animés ou inanimés qui capture l'attention humaine. Ces objets sont dotés de caractéristiques captivantes comme la nouveauté, la rareté et la distinctive par rapport à l'entourage. La figure 18 montre quelques exemples de tels objets. La problématique consiste donc à extraire de l'image les informations ou propriétés nécessaires pour distinguer l'objet du fond.



figure 18: Quelque exemple d'objets d'intérêt

Il existe deux grandes familles de repères visuels exploitées pour la détection d'objets saillants: les indices intrinsèques et les indices extrinsèques. Les indices intrinsèques font référence à des caractéristiques comme la couleur, la texture et les contours qui sont extraits sur la même image à traiter fin d'estimer l'objet saillant et supprimer les distracteurs. Cependant, certaines approches de détection d'objets saillants prétendent que certains objets possèdent des attributs visuels communs avec le fond ou les distracteurs. De ce fait, les indices intrinsèques s'avèrent insuffisantes pour distinguer l'objet du fond. Ils incorporent alors les indices extrinsèques qui sont extraites de sources d'informations externes à l'image. Nous pouvons citer comme exemple l'annotation de l'image par un utilisateur [21,50] et l'étude statistique sur des images similaires [51].

II.5. En quoi consiste la détection d'objets saillants ?

Détection d'objets saillants ou segmentation d'objets saillants est communément interprété en vision par ordinateur comme processus qui comprend deux étapes :

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

Cependant, les modèles sont rarement distingués explicitement entre ces deux étapes (avec quelques exceptions telles que Refs. [52-54]). Suivant les travaux fondateurs d'Itti et al. [55] et Liu et al. [56], modèles adoptent le concept de saillance pour simultanément effectuer les deux étapes ensemble. En témoigne le fait que ces étapes n'ont pas été séparées évaluées. En outre, la plupart des scores basés sur la zone ont été utilisés pour l'évaluation du modèle (p. ex. précision– rappel). La première étape ne doit pas nécessairement être limitée à un seul objet, La majorité des cependant, les modèles tentent de segmenter les objets, bien que leurs cartes de prédiction puissent être utilisées pour trouver plusieurs objets dans une scène. La deuxième étape tombe dans le domaine des problèmes de segmentation classiques en vision par ordinateur, mais avec la différence qu'ici, la précision n'est déterminée que par l'objet le plus saillant.



figure 19:Exemple d'un objet saillant

En général, il est convenu que pour une bonne saillance détection un modèle doit répondre au moins aux critères suivants trois critères:

- 1) bonne détection : la probabilité de régions réelles saillantes manquantes et de fausses marques le fond en tant que région saillante doit être faible,
- 2) haute résolution : les cartes de saillance devraient avoir haute ou pleine résolution pour localiser avec précision le saillant objets et conserver les informations originales de l'image;
- 3) efficacité de calcul : en tant que frontaux à d'autres processus complexes, ces modèles devraient détecter les rapidement.

II.6.État de l'art sur la détection d'objets saillants

II.6.1. Méthodes basées sur l'estimation du contraste local:

La notion de contraste global d'une image n'est pas forcément adaptée à la photographie, ce qui explique qu'on lui préfère souvent une définition locale du contraste qui correspond à la perception des contours des éléments d'une image. On parle souvent de piqué en photographie pour désigner le contraste des contours.

La loi de Weber-Fechner définit le contraste C_w comme:

$$C_w = \frac{L_{zone} - L_{fond}}{L_{fond}} \dots \dots \dots II.1$$

L_{zone} et L_{fond} désignent respectivement la luminance de l'objet et du fond directement aux abords de l'objet. Ce contraste mesure la perception que l'on a d'un objet de luminance L placé sur un fond de luminance L_{fond} . Quand le fond est plus clair que l'objet, C_w est négatif et varie entre 0 et -1. Quand le fond est plus foncé, C_w est positif et varie de 0 à, potentiellement, de grands nombres.

Une définition simplifiée est donnée par :

$$C_k = \frac{L_{zone}}{L_{fond}} [57] \dots \dots \dots II.2.$$

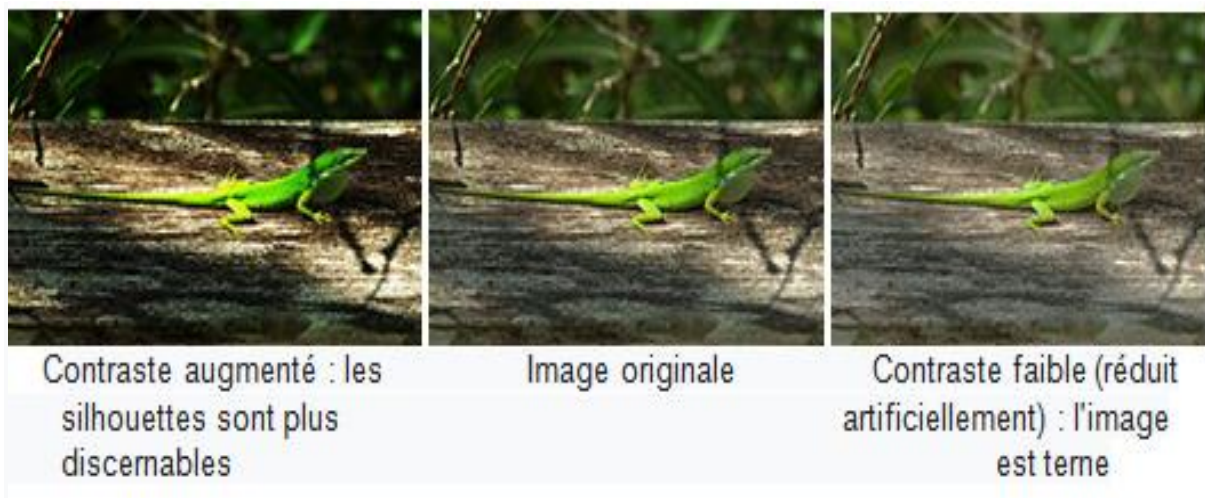


figure 20: exemple sur l'estimation du contraste local

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. Rappelons que les cellules bipolaires et ganglionnaires de la rétine peuvent former deux structures logiques comprenant les parties on-center et off-center .

Ces deux structures peuvent être utilisées pour extraire les parties les plus informatives d'une scène visuelle. L'une des méthodes les plus influentes est celle de Laurent Itti [4]. Dans cette approche, la saillance visuelle pour un emplacement est quantifiée comme sa différence à partir des emplacements voisins en considérant de multiples caractéristiques et des échelles multiples. L'architecture de ce modèle est basée sur les étapes suivantes. L'image source est décomposée en différents canaux perceptuels (dans la version de base de l'algorithme: intensité, couleur et orientation). Puis, une représentation multi-échelles pyramidale de l'image est construite à partir de chaque canal des caractéristiques utilisées.

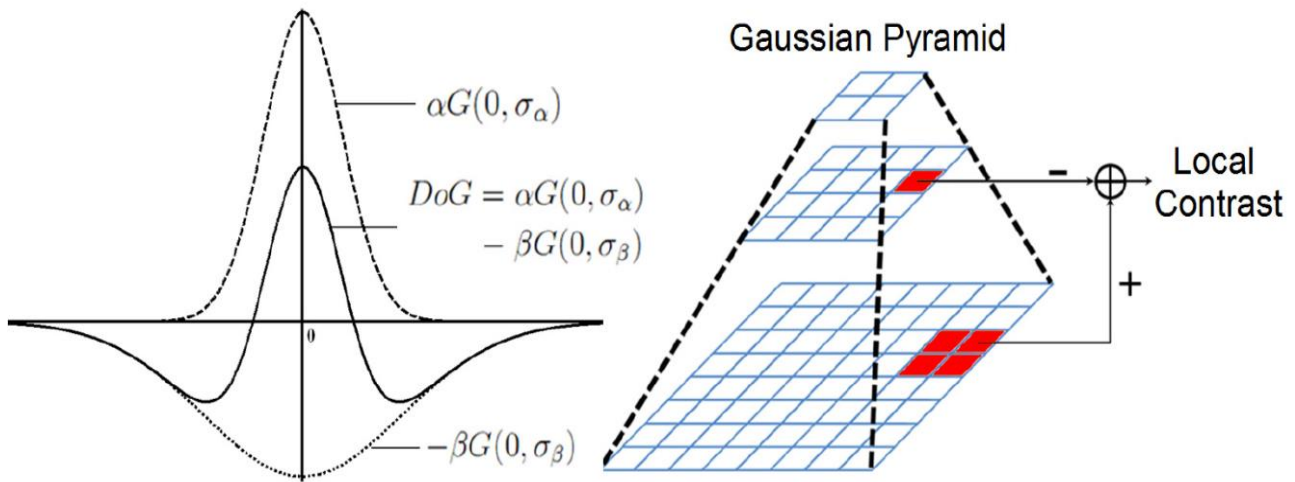


Figure 21 :Structure pyramidale dont chaque niveau correspond à une échelle de représentation. La différence de gaussiennes est utilisée pour estimer le contraste centre/périphérie à travers la pyramide [58].

Une opération de filtrage centre-périphérie est alors effectuée par soustraction de certains niveaux des pyramides multi-échelles afin d'obtenir différentes cartes de caractéristiques (cf. figure 21). Celles-ci sont ensuite normalisées par un opérateur permettant de renforcer les cartes ne contenant qu'un petit nombre de pics d'activité (zones saillantes) puis sommées afin d'obtenir 3 cartes de singularité (intensité, couleur et orientation). Ces cartes sont également normalisées, puis sommées afin d'obtenir une carte de saillance.

Le contraste local peut donc bien simuler le fonctionnement neurobiologique de la vision humaine. Par conséquent, beaucoup de modèles de saillance ont adopté une architecture similaire. Par exemple, Hu et al. [59] utilisent la différence de texture entre une région et ses régions voisines comme critère pour estimer le contraste local. Gao et al. [22] 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. Le modèle d'Olivier Le Meur [60] pousse l'inspiration biologique bien plus loin que le modèle original. Il modélise finement de nombreux phénomènes psycho-visuels: fonctions de sensibilité aux contrastes, décomposition en canaux perceptuels complexes, phénomène de masquage visuel, interaction centre/ périphérie, décomposition du flux perceptuel, etc. Cette approche montre qu'il est possible d'obtenir une modélisation très fine avec une approche plutôt macroscopique (le comportement individuel des neurones du cortex visuel n'étant pas du tout abordé). Kadir et al. [61] ont étudié la relation entre la saillance, l'échelle de présentation d'une région et ses caractéristiques. Ils proposent un modèle multi échelles basée sur l'entropie locale. Dans le même contexte, Les auteurs dans [62,63,64] utilisent la distance de χ^2 entre les histogrammes de couleur et de texture entre une région (ou bloc) et les régions (ou blocs) voisines pour l'estimation de la saillance locale.

II.6.2. Méthodes basées sur l'estimation du contraste global :

Le contraste global d'une image, souvent utilisé avec des mires de test pour déterminer la [fonction de transfert de modulation](#) d'un système optique est défini par Michelson par

$$C_m = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

Où L désigne la [luminance](#) et I l'[intensité lumineuse](#).

Ce contraste de Michelson est compris entre 0 et 1. Sa définition est analogue à celle de l'indice de [modulation](#) d'un signal modulé en amplitude[57].

Ces méthodes mesurent la rareté et l'unicité d'une région ou d'un bloc par rapport à l'ensemble de l'image. Plusieurs types d'approches ont été proposé 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 a 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 et l'utilisation du domaine fréquentiel ont été largement utilisés dans ce contexte.

II.6.2.1. Méthodes probabilistes:

Dans [65,66], les pixels d'une image I sont modélisés par des mixtures de gaussiennes (GMM) $\{w_c, \mu_c, \Sigma_c\}_{c=1}^C$, avec $\{w_c, \mu_c, \Sigma_c\}$ représente le poids, la moyenne colorimétrique et la matrice de covariance de la composante c respectivement. Chaque pixel x est affecté à une composante c avec la probabilité:

$$P(c|I_x) = \frac{w_c \mathcal{N}(I_x | \mu_c, \Sigma_c)}{\sum_c w_c \mathcal{N}(I_x | \mu_c, \Sigma_c)} \quad (2.1)$$

La variance spatiale horizontale de la composante c est définie comme suit:

$$V_h(c) = \frac{1}{|P|_c} \sum_x P(c|I_x) \|x_h - M_h(c)\|^2 \quad (2.2)$$

avec x_h représente les coordonnées dans l'axe des abscisses du pixel x et $|P|_c = \sum_x P(c|I_x)$.

Nous notons que la variance spatiale verticale $V_v(c)$ est définie d'une manière similaire. La distribution colorimétrique spatiale d'une composante c est alors définie comme suit:

$$DCS(c) = (1 - V(c)).(1 - D(c))$$

avec $V(c)$ la variance spatiale de la composante c , définie comme la somme des variances spatiales horizontales et verticales ($V(c) = V_h(c) + V_v(c)$) et $D(c)$ correspond à un paramètre de normalisation pondéré par la distance du pixel au centre de l'image d_x calculé comme suit:

$$D(c) = \sum_x P(c|I_x) \cdot d_x \quad (2.5)$$

La saillance finale d'un pixel x est alors calculée par rapport à toutes les composantes comme suit:

$$s(x) = \sum_c P(c|I_x).DCS(c) \quad (2.6)$$

II.6.2.2.Méthodes graphiques:

II.6.2.2.1Méthode graphique basée sur les marches aléatoires markoviennes:

Comme réponse aux limites des méthodes probabilistes, Harel et al. [67] 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 blocs 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:

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

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. Les poids entre les nœuds sont par la suite définis par la mesure de similarité et la distance spatiale comme suit:

$$w_k((i, j), (p, q)) = d_k((i, j) \parallel (p, q)) \cdot \exp\left(-\frac{(i-p)^2 + (j-q)^2}{2\sigma^2}\right) \quad (2.12)$$

avec σ est un paramètre de régularisation ajusté approximativement à un dixième ou un cinquième de la largeur de l'image. Une chaîne de Markov est alors appliquée sur le graphe où les nœuds correspondent à des états et les arcs à des transitions. Un nœud irrégulier possède des valeurs de similarité basse par rapport aux autres nœuds, il est donc moins visité dans la chaîne de Markov. Ce nœud possède une valeur de saillance élevée. Nous notons que dans les cas d'images avec plusieurs objets saillants et plusieurs distracteurs, les parties qui correspondent aux régions saillantes peuvent être visitées aussi fréquemment que les distracteurs. La performance de l'estimation de la saillance est alors considérablement affectée.

II.6.2.2.2.Méthode graphique basée sur la distance géodésique:

Dans [68], 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 a 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 .

$$\mathcal{E} = \{(r_i, r_j) | r_i \text{ est adjacent à } r_j\} \cup \{(r_i, B) | r_i \text{ est sur le contour de l'image}\} \quad (2.16)$$

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:

$$d_{geo}(r_i, r_k) = \min_{r_{i1}=r_i \dots r_{in}=r_k} \sum_{j=1}^{n-1} D_r(r_{ij}, r_{i,j+1}) ; (r_{ij}, r_{i,j+1}) \in \mathcal{E} \quad (2.17)$$

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

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

II.6.3. Méthodes basées sur la distance colorimétrique :

D'autres travaux [69,70] 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. [71] 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:

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

avec $D_r(\dots)$ 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:

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

avec $f(c_{k,i})$ la probabilité de la i ème couleur $c_{k,i}$ parmi toutes les n_k couleurs de la k ème région r_k , $k = \{1,2\}$. Cette probabilité est obtenue à partir de l'histogramme normalisé de r_k . $D(\dots)$ est la distance de couleur dans l'espace CIE Lab. Afin de favoriser la similarité entre les régions proches, les auteurs incorporent dans le processus d'estimation de la saillance la distance spatiale entre comme suit:

$$S(r_k) = \sum_{r_k \neq r_i} \exp(-D_s(r_k, r_i)/\sigma_s^2) w(r_i) D_r(r_k, r_i) \quad (2.15)$$

avec $D_s(r_k, r_i)$ est la distance spatiale euclidienne entre les centroides des régions r_k et r_i et σ_s un paramètre qui contrôle la contribution de la pondération spatiale. Plus la valeur de σ_s est importante, plus l'effet de la pondération spatiale est réduit et plus le contraste avec les régions éloignées contribuent dans l'estimation de la saillance.

II.6.4.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 [72] 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-spectre $\mathcal{L}(f)$ est ensuite calculé à partir de l'amplitude.

$$\mathcal{A}(f) = \mathcal{R}(\mathcal{F}[I(x)]), \quad \mathcal{P}(f) = \varphi(\mathcal{F}[I(x)]), \quad \mathcal{L}(f) = \log(\mathcal{A}(f)) \quad (2.19)$$

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:

$$\mathcal{R}(f) = \mathcal{L}(f) - h_n(f) * \mathcal{L}(f) \quad (2.20)$$

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

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

avec $g(x)$ un filtre gaussien qui assure un meilleur effet visuel. Guo et al. [73] ont démontré que l'incorporation de la phase de la transformée de Fourier au lieu de l'amplitude améliore les résultats sur la carte de saillance.

Achanta et al. [18] 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:

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

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

Au sein des méthodes basées sur l'estimation du contraste locale, l'unicité d'une région/pixel par rapport à toute l'image n'est pas prise en compte et sa saillance n'est pas détectée si celle-ci n'est pas contrastée par rapport à son entourage. En outre, comme l'échelle de l'objet dans l'image est inconnue a priori, il est particulièrement difficile de déterminer la taille des régions centrales et les régions voisines à comparer. Dans ce contexte, souvent, les contours qui sont les parties les plus contrastées par rapport à leur entourage sont retournés. Ces méthodes connaissent donc un taux de faux négatifs assez élevé et ne sont pas adaptées à une détection efficace et entière de tout l'objet saillant dans sa globalité (cf. figure 22 (c)).

Les méthodes basées sur le contraste global sont par contre plus orientées pour la séparation de l'objet saillant entier du reste de l'image. Cependant, ce type de méthode affecte souvent des valeurs de saillance élevées pour certaines régions de l'arrière-plan, spécialement dans le cas d'images avec un fond complexe. En effet, comme ces méthodes ne sont pas adaptées pour l'encodage de la contiguïté spatiale, certaines régions de l'arrière plan ne sont pas efficacement écartées. Un exemple est montré sur la figure 22 (d)-première rangée, où certaines régions de l'arrière-plan sont retournées comme saillantes. Ces méthodes sont également particulièrement limitées dans le cas de la similarité entre les distributions

colorimétriques de l'objet saillant et de l'arrière-plan (cf. figure 22 (d)- deuxième et troisième rangées). En outre, comme ces méthodes se basent sur les statistiques globales de l'image, elles sont dépendantes de la surface occupée par l'objet saillant. Lorsque l'objet saillant occupe un espace majoritaire dans l'image, l'estimation de la saillance retourne un taux important de faux négatifs et l'objet saillant dans sa globalité est mal détecté (cf. figure 22 (d)-quatrième figure).

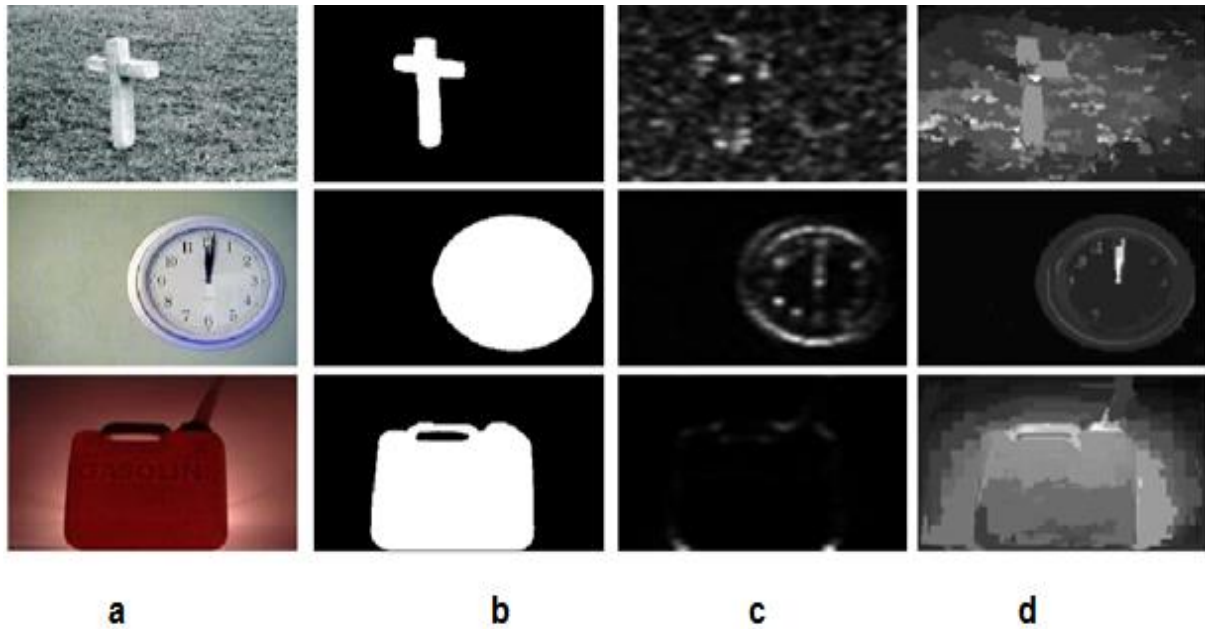


Figure 22 : Limite des méthodes basées sur l'estimation du contraste local et celles basées sur l'estimation du contraste global.

(a) Image en couleur, (b) vérité terrain, (c) carte de saillance retournée par une méthode basée sur l'estimation du contraste local [59] et (d) carte de saillance retournée par une méthode basée sur l'estimation du contraste global [18].

II.7. Conclusion:

Dans ce chapitre, on a présenté la notion de saillance et à quoi correspond la détection d'objets saillants puis on a fait une évaluation des différentes méthodes principales de détection d'objets saillants.

CHAPITRE III :

Application ,Résultat et discussion

III.1.Introduction :

Nous présentons dans ce chapitre quelques résultats de la méthode utilisée ,Nous détaillons dans un premier temps la base de données utilisée. Nous présentons par la suite quelques résultats expérimentaux en utilisant cette base .

Ensuite ,nous nous attarderons a la présentation de la méthode adoptée pour la détection des objets saillants nous passons à l'expérimentation et à la programmation ,des segmentation superpixelSLIC et saliencymap ,L'estimation de saillance peut être considérée comme une instance de segmentation d'image et qui concentre sur l'objet saillant et ignore les autres éléments d'arrière-plan , selon l'étude effectuée nous avons choisit l'algorithme développé par Chuan Yang et al pour la segmentation d'image ,Les superpixels sont de plus en plus utilisés dans les applications de vision par ordinateur. Cependant, il existe peu d'algorithmes qui produisent un nombre souhaité de superpixels réguliers et compacts avec une faible surcharge de calcul. Nous introduisons un nouvel algorithme appelé SLIC (Simple Linear Iterative Clustering).

III.2.L'Environnement de Développement que nous avons Utilisé

Pour l'implémentation nous avons choisi l'environnement Matlab, reconnu pour sa robustesse dans les calculs matriciels. La version de Matlab est R2016a, sur une machine caractéristique: Génuine Intel CPU et une mémoire RAM de 4 Go.

MATLAB est un langage initialement créé pour traiter des problèmes d'analyse numérique. Il est optimisé pour le calcul matriciel et convient donc parfaitement pour la manipulation et le traitement d'images numériques. Ces dernières sont en effet représentées par des tableaux à deux ou trois dimensions (2D ou 3D) .

III.3.Base de données utilisées:

Nous appliquons notre méthode sur les ensembles de données d'images disponibles au public MSRA-1000 pour évaluer ses performances. L'ensemble de données MSRA contient 1000 images couleur avec des segmentations précises des contours d'objet par pixel, qui sont sélectionnées à partir d'un ensemble de données de 5000 images
Voici quelques exemples d'images existantes dans la base de donnée



Figure 23: quelques exemple d'images de la base de données

III.4. Présentation de la Méthode adoptée et interprétation des résultats :

III.4.1. Evaluation de la pertinence des caractéristiques :

L'Evaluation de la pertinence des caractéristiques a comme objectif d'attribuer plus d'importance pour les caractéristique qui discriminent mieux l'objet saillant de l'arrière-plan. Pour une caractéristique donnée nous considérons ses histogramme normalisé dans l'intervalle $[0,1]$ notés p et q au sien du noyau de l'objet et l'arrière-plan respectivement . Nous supposons que chaque caractéristique est échantillonnée sur 256 bins et soient $p(i)$ et $q(i)$ les fréquences normalisées du i eme bin .le pouvoir discriminatif d'une caractéristique est représenté par le degré de chevauchement entre p et q .ceci peut être formalisé comme suit :

$$V = \frac{\text{var}(F; \frac{p+q}{2})}{\text{var}(F;p) + \text{var}(F;q)} \dots \dots \text{III.1}$$

Avec $\text{var}(F;p)$, $\text{var}(F;q)$ et $\text{var}(F; \frac{p+q}{2})$ représentent la variance de la fonction F obtenue par rapport aux distribution p , q et $\frac{p+q}{2}$.

La variance de F par rapport a une distribution R est calculée comme suit :

$$\text{var}(F, R) = \sum_1^{256} R(i) * F(i)^2 - [\sum_1^{256} R(i) * F(i)]^2 \dots \dots \text{III.2}$$

La valeur de F pour le i eme bin est calculée comme suit:

$$F(i) = \log\left(\frac{\max(p(i), \epsilon)}{\max(q(i), \epsilon)}\right) \dots \dots \text{III.3}$$

Avec une petite valeur qui évite les fréquences nulles. Cette valeur est ajustée empiriquement a 0.001. Nous notons que par un tel processus ,il est facile de constater que V soit défini dans l'intervalle $[0.5 ; +\infty]$ Afin de normaliser l'évaluation de la pertinence dans l'intervalle $[0,1]$, nous utilisons la fonction suivant :

$$\epsilon = 1 - \exp(-\delta |V - 0.5|) \dots \dots \text{III.4}$$

Avec δ un facteur qui contrôle la sensibilité du calcul de la pertinence. Augmenter la valeur de ce facteur augmenterait le nombre de caractéristique considérées pertinentes, et vice versa ce vecteur est ajusté empiriquement à 0.2.

III.4.2 Segmentation par diffusion graphique :

la méthode qu'on a utilisé est la méthode graphique basée sur les marches aléatoires markoviennes qui 'une des Méthodes basées sur les graphes qui ont récemment émergé comme un outil efficace et fort pour la détection d'objets saillants et plus précisément est une méthode basées sur une seule résolution qui se base 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.

La segmentation superpixellique consiste à travailler non pas sur des pixels mais sur des groupes de pixels homogènes appelés super-pixels qu'ils décrits par 'algorithme SLIC qui est un K-Means adapté à la génération de superpixels

L'algorithme SLIC permet de segmenter une image en K régions appelées 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. Plutôt que de faire une simple décimation sur l'image pour diminuer la quantité d'information, segmenter en superpixels permet d'avoir un ensemble de régions d'intérêts à traiter, sans pour autant diminuer la quantité d'information brute dans l'image

Marches aléatoires markoviennes:

Dans [54], 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. Le poids reliant deux nœud i et j est défini comme suit:

$$w_{i,j} = e^{-\frac{\|x_i - x_j\|}{\sigma^2}}; \quad i, j \in V \dots \dots \dots \text{ III.5}$$

avec x_i et x_j les moyennes des nœuds i et j respectivement dans l'espace de couleur CIE Lab et σ une constante qui contrôle la stabilité du poids. Les nœuds sont ensuite réorganisés de façon que les t premiers nœuds soient transitoires et les r derniers soient absorbants. Ils définissent alors la matrice d'affinité A comme suit:

$$A: \begin{cases} w_{ij} \in N(i), & 1 \leq i \leq t \\ 1 \text{ si } i = j & \dots \dots \dots \text{ III.6} \\ 0 \text{ sinon} \end{cases}$$

avec $N(i)$ l'ensemble des nœuds connectés à i . La matrice de transition P est alors calculée comme suit:

$$P = D^{-1} \times A \dots \dots \text{III. 7}$$

avec D la matrice des degrés qui représente la somme des poids connectés à chaque nœud $D = \text{diag}(\sum_j a_{ij}), j \in t + r$. 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. La matrice P possède la forme canonique suivante $P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix}$ avec $Q \in [0,1]^{t \times t}$ la matrice contenant les probabilités de transition entre chaque paire de nœuds transitoires et $R \in [0,1]^{t \times r}$ 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 $r \times t$ et 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. La valeur de saillance de tous les nœuds peut donc être estimée comme suit:

$$\bar{y} = N \times c \dots \dots \dots \text{III. 8}$$

où c est un vecteur de taille t dont les éléments sont tous égaux à 1 et \bar{y} une valeur normalisée dans $[0,1]$. 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. Les auteurs se basent 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 remédier à cela, les auteurs calculent un score pour estimer le degré de ce contraste comme suit:

$$\text{score} = \sum_{b=1}^{10} g(b) \times \min(b, (11 - b)) \dots \dots \text{III. 9}$$

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. 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:

$$\pi_i = \frac{\sum_j a_{ij}}{\sum_{ij} a_{ij}} \dots \dots \quad \text{III. 10}$$

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, les auteurs utilisent le temps récurrent moyen $h_j = 1/\pi_j$ pour chaque nœud j afin de pondérer chaque élément n_{ij} correspondant dans la matrice fondamentale N .



Figure 24 : Des exemples illustrant les avantages du traitement de mise à jour de la saillance [54]. (a) lamage en couleur, (b) carte de saillance générée sans le traitement de mise à jour, (c) carte de saillance générée avec le traitement de mise à jour.

Ils substituent ensuite le vecteur c dans l'équation III.8 par le vecteur u défini comme suit: $u_j = h_j / \sum_k h_k$ avec k et j des indexes des nœuds transitoires sur le graphe. Cela permet d'atténuer le niveau de saillance de ces zones au centre de l'image. Un inconvénient majeure de cette méthode réside dans le module de mise à jour qui peut affecter considérablement le taux de vrais positifs au centre de l'image dans le cas d'objets saillants définis sur une structure largement homogène avec des valeurs de saillance moyennes.

III.4.3. Evaluation des performances :

III.4.3.1 Classification des pixels :

Afin d'analyser la qualité de la détection, l'image de différence obtenue après traitement, est comparé avec l'image de vérité terrain correspondante. Les pixels peuvent être classifiés en 4 groupes différents :

- 1) Vrai Positif (VP) : représente le nombre des pixels du premier plan correctement détectés par l'algorithme.
- 2) Faux Positif (FP) : représente le nombre des pixels qui sont incorrectement classifiés comme objet de premier plan.
- 3) Vrai Négatif (VN) : représente le nombre des pixels appartenant à l'arrière plan et qui sont correctement détectés comme arrière-plan de la scène par l'algorithme.
- 4) Faux Négatif (FN) : représente le nombre des pixels appartenant au premier plan qui sont mal classés en tant que pixels appartenant à l'arrière-plan.

III.4.3.2. Mesures d'évaluation comparatives :

III.4.3.2.1 .La Sensibilité (Recall) :

La sensibilité (taux de vrai positif) est définie comme le rapport du nombre des vrais positifs sur le nombre total des éléments qui appartiennent aux objets du premier plan (c'est-à-dire la somme des deux : vrai positif et faux négatif) :

$$\text{sensibilité} = \frac{vp}{vp+FN} \dots \dots \dots \text{III.11}$$

En d'autres termes il se traduit comme suit :

La sensibilité est généralement désignée par le terme de rappel

III.4.3.2.2 .La précision :

La précision (valeur prédictive positive) peut être considérée comme une mesure de d'exactitude ou de fidélité et est évaluée par la division du nombre d'éléments (objets du premier plan) correctement détectés sur le nombre totale des pixels classifiés comme premier plan par l'algorithme. En réalité la précision évalue la décision de l'algorithme à définir un pixel soit appartenant au premier plan (objet) ou a l'arrière plan et a quel point cette décision est fiable :

$$\text{Précision} = \frac{vp}{vp+fp} \dots\dots\dots \text{III.12}$$

III.4.3.2.3 La spécificité :

La spécificité (taux des vrais négatifs) concerne la capacité de l'algorithme a détecter correctement les pixels de l'arrière-plan, la spécificité est le rapport du nombre des vrais négatif sur le nombre totale des éléments qui appartiennent a l'arrière-plan, elle est exprimée comme suit :

$$\text{spécificité} = \frac{VN}{VN+FP} \dots\dots\dots \text{III.13}$$

III.4.3.2.4. Accuracy:

$$\text{accuracy} = (tp + tn)/vp + vn + 2fn \dots\dots \text{III.14}$$

III.4.3.2.5.F_measure :

La F_measure permet d'agréger le rappel et la précision dans une mesure unique ,elle est définie comme la moyenne harmonique pondérée du rappel et de la précision

$$\text{F_measure} = \frac{2*(\text{precision}*\text{recall})}{(\text{precision}+\text{recall})} \dots\dots\dots \text{III.15}$$

III.4.3.2.6.Gmean:

$$\text{gmean} = \sqrt{\text{spécificité} * \text{Sensibilité}} \dots\dots\dots \text{III.16}$$

III.4.3.2.7.Xor:

$$xor = \frac{fp+fn}{vp+fn} \dots \text{III.17}$$

Tableau de comparaison entre les cartes de saillance et masque de quelques images:

Image	Res							
Carte saillante /masque	accuracy	sensitivité	specificité	xor	precision	rappel	F_measure	gmean
	.9462	0.9907	0.9366	0.3040	0.7707	0.9907	0.8670	0.9633

on a fait ici une comparaison entre la carte saillances et masque binaire obtenue sur tout l'ensemble de la base de donnée MSRA-1000 et on a obtenu les valeurs indiquées dans ce tableau

on a la valeur de sensibilité (rappel)=99.07% donc le taux de vrai positif >50%

laspecificité =93.66% donc le taux de vrai négatif >50%

laprecision =77.07% donc la valeur prédictive positive >50%

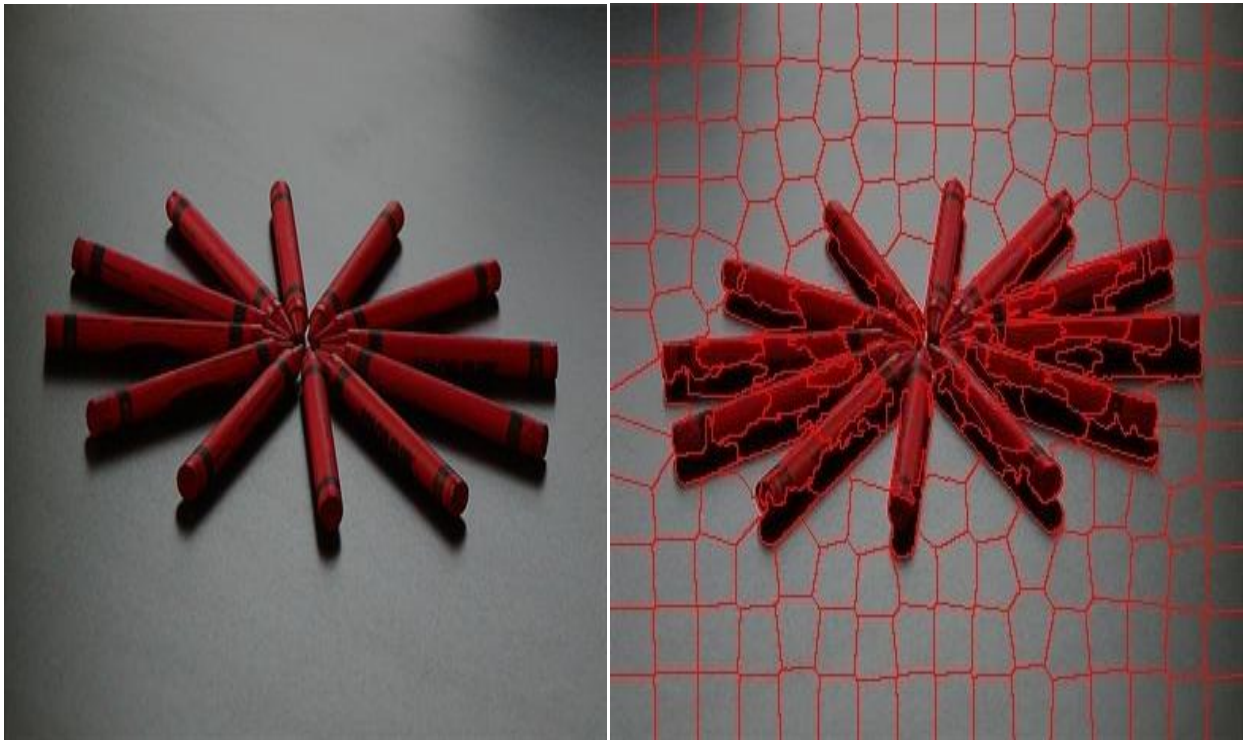
accuracy=94.62% >50%

xor =30.40% <50% c'est a dire que la valeur de faux positif et de faux négatif est plus petite

ce qui implique qu'on a une bonne similarité entre la carte saillance et le masque binaire donc un bonne détection d'objet saillant.

III.5.Résultats expérimentaux:

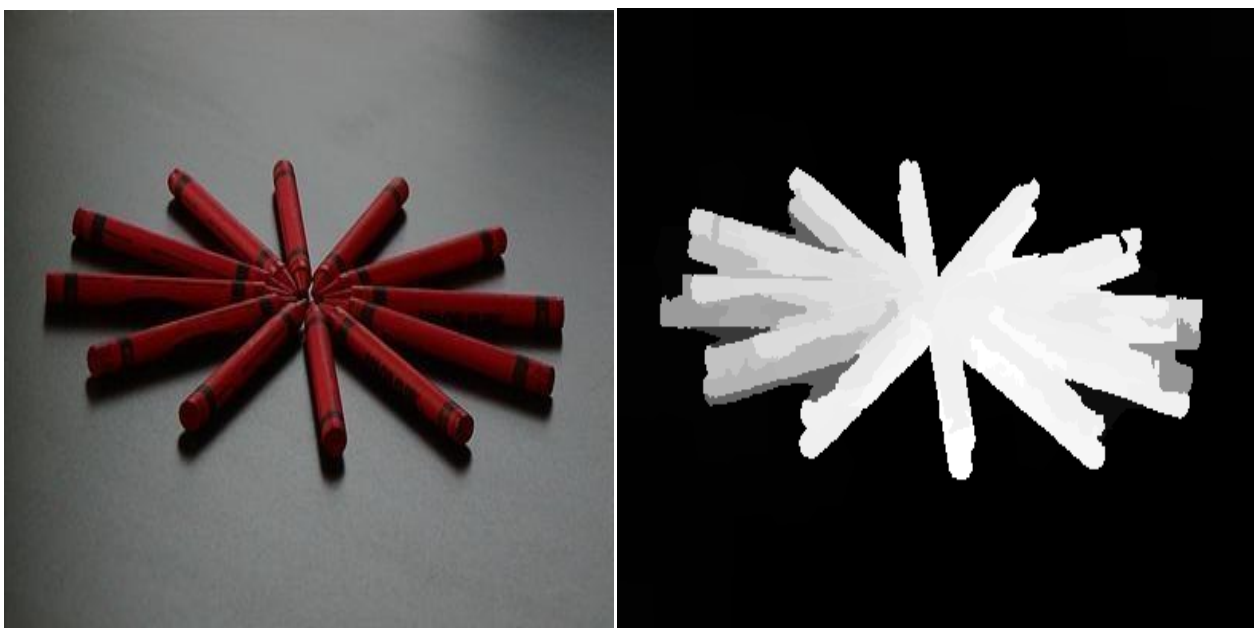
III.5.1.Présentation des résultats:



A

B

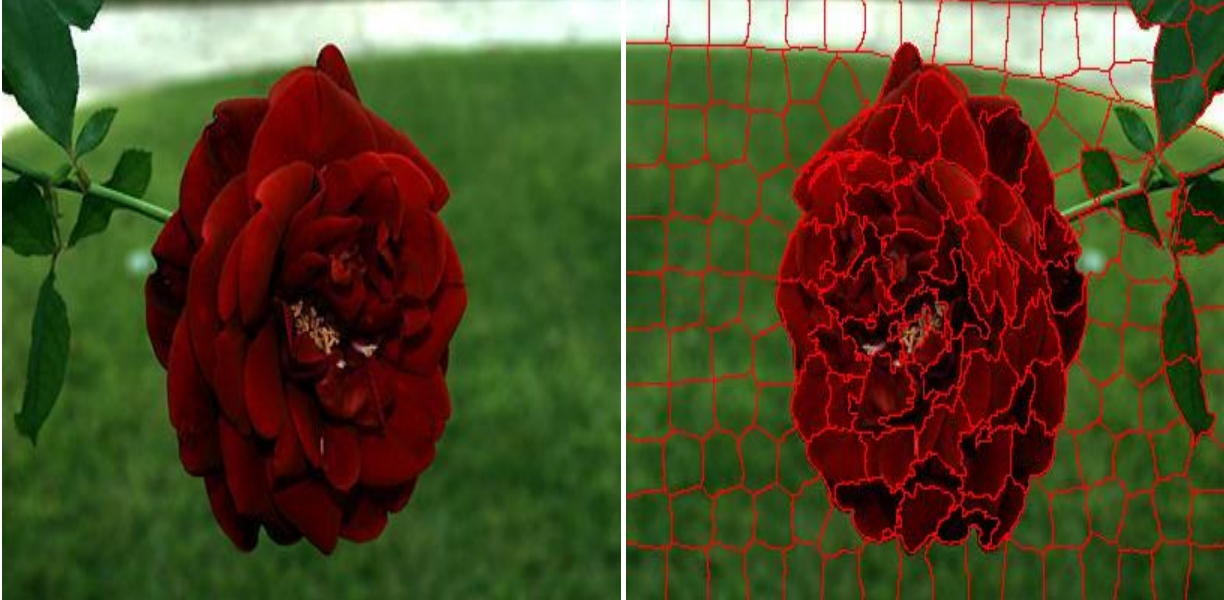
Figure 25 : B:résultat de segmentation en superpixel de l'imageA



C

D

Figure 26: la carte de saillance obtenue de l'image C



E

F

Figure 27 :F: résultat de l'image E en superpixel



G

H

Figure 28 :H: la carte de saillance obtenue de l'image G



I



J

Figure 29 : J:résultat de la segmentation en superpixel de l'image I



K



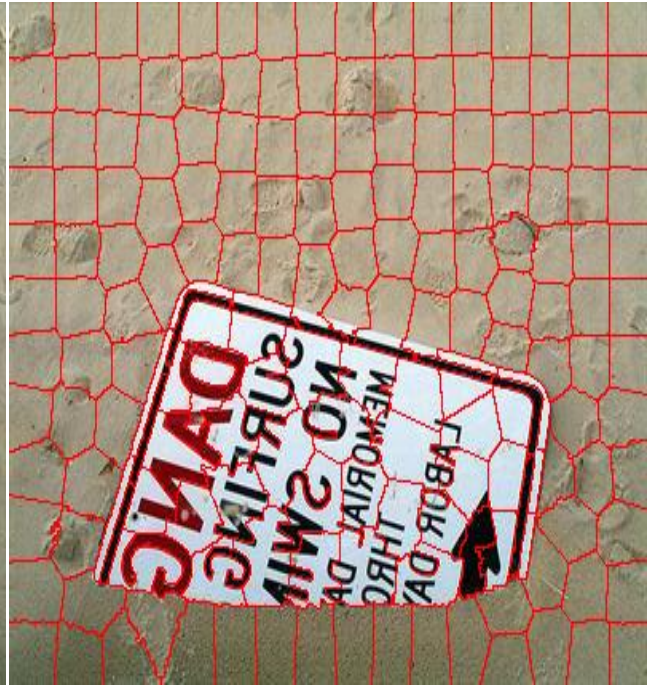
L

TV

Figure 30 : la carte de saillance obtenue de l'image K



M



N

Figure 31:N:résultat de la segmentation superpixel de limageM



O



P

Figure 32 :la carte saillance obtenue de l'image O



Q



R

Figure 33 : R:résultat obtenue de la segmentation en superpixel de 'l'image Q



S

T

Figure 34 :T:la carte de saillance obtenue de l'image S

III.6.Conclusion :

Dans ce chapitre, nous avons présenté les résultats obtenus de la détection des objets dans une image à partir d'un algorithme exploité, se combinaison avec une base de données utilisées permettant résultats expérimentaux afin de valider notre méthode proposée ainsi que la présentation des résultats et analyse et discussion avec des remarques sur notre travail que toutes les images avaient donnés une bonne détection.

Le travail que nous avons effectué nous a permis de réaliser à l'aide de l'environnement MATLAB, l'objectif de notre mémoire.

conclusion générale :

Dans ce mémoire, nous nous sommes intéressés à la détection la détection d'objets saillants dans des scènes complexe.

Le travail présenté dans ce mémoire consiste à réaliser un système permettant la détection des objets saillants dans des images en utilisant la diffusion graphique a l'aide de l'environnement MATLAB ;Aussi on a évalué la pertinence des caractéristiques et cela en passant par plusieurs étapes :

la première étape ,on a d'abord segmenter les images en superpixels avec la méthode SLIC qui génère des superpixels en regroupant les pixels en fonction de leur similarité de couleur et la proximité dans le plan de l'image; Ensuite ,on a généré les cartes de saillance ;Finalement , on a comparé les images résultantes de la carte saillance avec la vérité terrain(maque binaire)puis on a trouvé des résultats qui ont montré l'efficacité de cette méthode.

Les résultats obtenus sont visuellement très clairs, précis et des valeurs tres bonnes qui montrent d'une part que l'approche proposée dans ce travail permet de détecter avec succès les objets saillants , alors nous suggérons de continuer la recherche dans ce sens pour progresser encore plus les résultats et donc plusieurs améliorations peuvent être apportées ;

Dans le futur nous espérons implémenter sous java notre travail sous forme de logiciel ou application du détection d'objets ;

Bibliographie :

- [1] : Arnaud CAPRI « 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 » thèse présentée à l'université d'Orléans pour obtenir le grade de : Docteur de l'université d'Orléans 2007.
- [2] : MEMOIRE Une Approche semi-automatique pour l'indexation de documents anciens Présenté en vue de l'obtention du diplôme de MAGISTER Intitulé
- [3]:http://www.mit-university.net/index.php?view=entry&id=2%3Ala-notion-depixel&option=com_lyftenbloggie&Itemid=123 ,Août 2010.
- [4]: <http://www.commentcamarche.net/contents/video/affich.php3>
- [5] K.Tabari,S.Tagma «Compression d'images animées à très faible débit par la géométrie des fractales» PFE USTHB 1996
- [6]: <http://www.crdp.ac-grenoble.fr/image/general/formats.htm>
- [7] :http://www.tsi.enst.fr/pages/enseignement/ressources/beti/RVB_ou_LAB/html/colorspace.html
- [8] :M .T.Chikh , amélioration des images par un modèle de réseau de neurones (comparaison avec les filtres de base), université abou-bakrbelkaid tlemcen.2011
- [9]: J.F. Canny. Finding edges and lines in images, M.I.T. Artificial Intelligence Laboratory, Cambridge, MA, TR-720, 1983.
- [10] : S. Beucher et F. Meyer. The morphological approach to segmentation:

The watershed transformation. *Mathematical Morphology in Image Processing*, pp. 433-481, 1993

[11] : S.L. Horowitz et T. Pavlidis. Picture segmentation by a directed split and merge procedure. In *Computer Methods in Images Analysis*, pp. 101-11, 1977.

[12]: R.M. Haralick et L.G. Shapiro. *Computer and Robot Vision*, volume 1. Addison-Wesley, 1992.

[13] : J. G. Aviña Cervantès. Navigation visuelle d'un robot mobile dans un environnement d'extérieur semi-structuré. Thèse de Doctorat, Institut National Polytechnique de Toulouse, Février 2005.

[14]: J. A. Hardigan. *Clustering Algorithms*. John Wiley & Sons, New York, 1975.

[15]: J. McQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281-297, 1967.

[16]: <https://www.tu-chemnitz.de/etit/proaut/en/research/superpixel.html>.

[17] : A. Borji and L. Itti. Exploiting Local and Global Patch Rarities for Saliency Detection. *IEEE Conf. on Computer Vision and Pattern Recognition*, 478- 485, 2012.

[18]: 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

[19] : A. Borji and L. Itti. State-of-the-art in Visual Attention Modeling. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 35(1):185-207, 2013.

[20]: X. Hou and L. Zhang. Saliency Detection: A Spectral Residual Approach. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007.

[21] : A. Borji, M.-M. Cheng, H. Jiang and J. Li. Salient Object Detection: A Survey. [online] Available: <http://arxiv.org/abs/1411.5878>, 2014.

[22] : D. Gao, V. Mahadevan, and N. Vasconcelos. The discriminant center surround hypothesis for bottom-up saliency. In: *Advances in Neural Information Processing Systems*, 2009.

[23] : <https://imotions.com/blog/saliency-detection/>

[24] : Itti, L.; Koch, C.; Niebur, E. A model of saliencybased visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 20, No. 11, 1254–1259, 1998

[25] : Treisman, A. M.; Gelade, G. A feature-integration theory of attention. *Cognitive Psychology* Vol. 12, No. 1, 97–136, 1980.

[26] : Wolfe, J.M.; Cave, K.R.; Franzel, S.L. Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance* Vol. 15, No. 3, 419–433, 1989.

[27] : Koch, C.; Ullman, S. Shifts in selective visual attention: Towards the underlying neural circuitry. In: *Matters of Intelligence. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science)*, Vol. 188. Vaina, L. M. Ed. Springer Dordrecht, 115–141, 1987.

[28] : Parkhurst, D.; Law, K.; Niebur, E. Modeling the role of salience in the allocation of overt visual attention. *Vision Research* Vol. 42, No. 1, 107–123, 2002.

- [29] : Bruce, N. D. B.; Tsotsos, J. K. Saliency based on information maximization. In: Proceedings of the 18th International Conference on Neural Information Processing Systems, 155–162, 2005.
- [30] : Liu, T.; Sun, J.; Zheng, N.; Tang, X.; Shum, H.-Y. Learning to detect a salient object. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1–8, 2007.
- [31] : Liu, T.; Yuan, Z. J.; Sun, J.; Wang, J. D.; Zheng, N. N.; Tang, X. O.; Shum, H.-Y. Learning to detect a salient object. IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 33, No. 2, 353–367, 2011.
- [32] : Achanta, R.; Estrada, F.; Wils, P.; Sussstrunk, S. Salient region detection and segmentation. In: Computer Vision Systems. Lecture Notes in Computer Science, Vol. 5008. Gasteratos, A.; Vincze, M.; Tsotsos, J. K. Eds. Springer Berlin Heidelberg, 66–75, 2008.
- [33] : Ma, Y.-F.; Zhang, H.-J. Contrast-based image attention analysis by using fuzzy growing. In: Proceedings of the 11th ACM International Conference on Multimedia, 374–381, 2003.
- [34] : Liu, F.; Gleicher, M. Region enhanced scale-invariant saliency detection. In: Proceedings of the IEEE International Conference on Multimedia and Expo, 1477–1480, 2006.
- [35] : Walther, D.; Koch, C. Modeling attention to salient proto-objects. Neural Networks Vol. 19, No. 9, 1395–1407, 2006.
- [36] : Arbel'aez, P.; Maire, M.; Fowlkes, C.; Malik, J. Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 33, No. 5, 898–916, 2011.
- [37] : Martin, D. R.; Fowlkes, C. C.; Malik, J. Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 26, No. 5, 530–549, 2004.
- [38] : Cheng, M.-M.; Zhang, Z.; Lin, W.-Y.; Torr, P. H. S. BING: Binarized normed gradients for objectness estimation at 300fps. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3286–3293, 2014.
- [39] : Alexe, B.; Deselaers, T.; Ferrari, V. What is an object? In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 73–80, 2010.
- [40] : Endres, I.; Hoiem, D. Category independent object proposals. In: Computer Vision – ECCV 2010. Lecture Notes in Computer Science, Vol. 6315 Daniilidis, K.; Maragos, P.; Paragios, N. Eds. Springer Berlin Heidelberg, 575–588, 2010.
- [41] : Judd, T.; Ehinger, K.; Durand, F.; Torralba, A. Learning to predict where humans look. In: Proceedings of the IEEE 12th International Conference on Computer Vision, 2106–2113, 2009.
- [42] : Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1–8, 2007.
- [43] : Borji, A.; Itti, L. Exploiting local and global patch rarities for saliency detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 478–485, 2012.
- [45] : Borji, A. Boosting bottom-up and top-down visual features for saliency estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 438–445, 2012.

- [46] : Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2001.
- [47] : Felzenszwalb, P. F.; Girshick, B.; McAllester, D.; Ramanan, D. Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 32, No. 9, 1627–1645, 2010.
- [48] : LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* Vol. 86, No. 11, 2278–2324, 1998
- [49] : Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440, 2015.
- [50] : M. Wang, J. Konrad, P. Ishwar, K. Jing, and H. A. Rowley. Image Saliency: From Intrinsic to Extrinsic Context. *IEEE Conf. Computer Vision and Pattern Recognition*, 417424, 2011.
- [51] : F. Huazhu, C. Xiaochun and T. Zhuowen. Cluster-based Co-saliency Detection. *IEEE Trans. on Image Processing*, 22(10):3766-3778, 2013.
- [52] : B. Alexe, T. Deselaers and V. Ferrari. Measuring the Objectness of Image Windows. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 34(11):2189-2202, 2012.
- [53] : M.E. Raichle, M.E.: The brain's dark energy. *Scientific American Magazine*, 44-49, 2010.
- [54] : C.Koch. Biophysics of Computation: Information Processing in Single Neurons, 1st edn. *Oxford University Press, New York*, 2004.
- [55] : J. Li, W. Gao. Visual Saliency Computation, a Machine Learning Perspective, *Lecture Notes in Computer Science, Springer*, 8408 2014, 2014.
- [56] : W. James. The principles of psychology. *Dover Publications, volume 1 edition*, 1890.
- [57] : https://fr.wikipedia.org/wiki/Contraste#Contraste_global_d'une_image
- [58] : L. Itti, C. Koch and E. Niebur. A Model of Saliency Based Visual Attention for Rapid Scene Analysis. *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 20(11):1254-1259, 1998.
- [59] : Y. Wei, F. Wen, W. Zhu and J. Sun. Geodesic saliency using background priors. *European Conf. on Computer Vision*, 29-42, 2012.
- [60] : O. Le Meur, P. Le Callet, D. Barba, D. Thoreau. A coherent computational approach to model bottom-up visual attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(5): 802–817, 2006.
- [61] : T. Kadir, M. Brady. Saliency, scale and image description. *International Journal of Computer Vision*, 45(2): 83-105, 2001.
- [62] : T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.Y. Shum, Learning to detect a salient object. *IEEE TPAMI*, 33(2): 353–367, 2011.
- [63] : T. Liu, J. Sun, N. Zheng, X. Tang and H.Y. Shum. Learning to detect a salient object. *in CVPR*, 1–8, 2007.
- [64] : D.R. Martin, C.C. Fowlkes, and J. Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE TPAMI*, 26(5): 530-549, 2004.
- [65] : T. Liu, J. Sun, N. Zheng, X. Tang and H.Y. Shum. Learning to detect a salient object. *in CVPR*, 1–8, 2007.

- [66] : M.M. Cheng, J. Warrell, W.Y. Lin, S. Zheng, V. Vineet, and N. Crook. Efficient salient region detection with soft image abstraction. in ICCV, 1529–1536, 2013.
- [67] :J. Harel, C. Koch and P. Perona. Graph-based visual saliency. *In: Advances in Neural Information Processing Systems*, 545-552, 2007.
- [68] : Y. Wei, F. Wen, W. Zhu and J. Sun. Geodesic saliency using background priors. *European Conf. on Computer Vision*, 29-42, 2012.
- [69] : 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.
- [70] : Y. Wei, F. Wen, W. Zhu and J. Sun. Geodesic saliency using background priors. *European Conf. on Computer Vision*, 29-42, 2012.
- [71] : 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.
- [72] : X. Hou and L. Zhang. Saliency Detection: A Spectral Residual Approach. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2007.
- [73]:C. Guo, Q. Ma, and L. Zhang. Spatio-Temporal Saliency Detection Using Phase Spectrum of Quaternion Fourier Transform. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008.