

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

Université Mouloud Mammeri de Tizi-Ouzou

Faculté de génie électrique et de l'informatique

Département d'informatique



MÉMOIRE

De fin d'études en vue de l'obtention du Diplôme Master

Option: Système d'informatique

THÈME

Détection d'objets saillants dans des images en couleurs

Présenté par :

M^{elle} Mansour Djamila

M^m Ferdji Hafidha

Dirigé par :

Mr Filali Idir (Maître Assistant)

Promotion 2014/2015

Remerciements

Nous tenons à remercier en premier lieu notre encadreur Mr Filali Idir enseignant chercheur à l'UMMTO de nous avoir fait confiance en nous proposant ce sujet. On veut le remercier pour sa disponibilité et du temps consacrer à ce travail.

A monsieur le président et à messieurs les membres de jury, pour l'honneur qu'ils nous ont fait, en acceptant d'évaluer notre travail

On remercie le vice recteur chargé des relations extérieures et manifestations scientifiques Pr. Ait-Aider Hacene et Mm Chamek Leila chef de service informatique au CHUTO, pour leur patience et leurs conseils avisés.

A Monsieur Rassoul Idir Maître de conférences au département d'informatique de l'UMMTO pour ses encouragements de reprendre les études et ses orientations précieuses.

A Monsieur Pr A.ZIRI Directeur général du CHU Tizi-Ouzou

Merci également à tous les enseignants du département informatique de l'université Mouloud Mammeri de Tizi-Ouzou qui nous ont aidés pendant nos années d'études et pour leurs conseils précieux.

On voudrait aussi remercier les collègues de l'université Mouloud Mammeri de Tizi-Ouzou en particulier ceux du vice rectorat chargé des relations extérieurs et manifestations scientifiques, ainsi que les collègues du CHUTO qui ont été compréhensifs et qui nous ont supportés et soutenus durant ce parcours.

Un grand merci aussi à tout notre entourage : père, mère, frères, sœurs, neveux, nièces, belles sœurs et beaux frères pour leurs soutiens permanents.

Enfin, c'est pour nous un plaisir autant qu'un devoir, d'exprimer notre gratitude et reconnaissance à toutes les personnes ayant contribué de près ou de loin à la réalisation de ce travail.



Dédicaces

Je dédie ce travail à :

*La mémoire de ma très chère sœur Malika, en témoignage de ma
profonde gratitude et amour, que son âme repose en paix.*

*Mes très chers parents pour leurs encouragements et leurs soutiens
durant chaque étape de ma vie et que dieu les protège.*

Mes chers frères et sœurs et leurs petites familles.

Mes neveux et nièces.

Toute ma famille paternelle et maternelle

Mes collègues de l'UMMTO en particulier l'équipe de l'auditorium.

Tous mes amis (es)

Mon binôme et toute sa famille

Toute la promotion 20014/20015 en particulier le groupe DEUA.

Djamila





Dédicaces

Je dédie ce travail à :

*Mes très chers parents pour leurs encouragements et leurs soutiens
durant chaque étape de ma vie et que dieu les protège.*

Mon mari Djilali et sa famille de m'avoir soutenu.

Mes chers frères et sœurs et leurs petites familles.

Mes neveux et nièces.

*Ma chef de service informatique Mme Chamek Leila pour sa
compréhension durant tout mon parcours.*

Tous mes collègues du service informatique du CNUSO en particulier :

Radia et Ossia

Tous mes amis (es)

Mon binôme et toute sa famille

Toute la promotion 20014/20015 en particulier le groupe DEUR.

Kafidha



Sommaire

<i>Introduction générale.....</i>	<i>01</i>
<i>Chapitre I : Généralités sur le traitement & segmentation d'image</i>	
<i>Partie I : Généralités sur le traitement d'images</i>	
<i>I. Introduction.....</i>	<i>05</i>
<i>II- Notions D'image</i>	<i>05</i>
<i>II.1. Image en niveau de gris</i>	<i>05</i>
<i>II.2. Image binaire.....</i>	<i>06</i>
<i>II.3. Image couleur.....</i>	<i>07</i>
<i>II.4. Caractéristique d'une image</i>	<i>07</i>
<i>III. Histogramme d'une image.....</i>	<i>09</i>
<i>IV. Les filtres.....</i>	<i>11</i>
<i>Conclusion...</i>	<i>12</i>
<i>Partie II : La segmentation d'images</i>	
<i>I- Introduction.....</i>	<i>14</i>
<i>II- Définitions.....</i>	<i>14</i>
<i>II- 1. Définition de la segmentation.....</i>	<i>14</i>
<i>II-2. Définition d'une région.....</i>	<i>15</i>
<i>III. Approches de la Segmentation.....</i>	<i>15</i>
<i>III- 1. Approche globale de la segmentation.</i>	<i>16</i>
<i>III-2. Approche locale de la segmentation..</i>	<i>16</i>
<i>III-3. Approche hybride de la segmentation.....</i>	<i>16</i>
<i>VI- Les Méthodes de segmentation.....</i>	<i>16</i>
<i>VI -1 Les histogrammes.....</i>	<i>16</i>
<i>VI -2 Méthodes k-means.....</i>	<i>19</i>
<i>VI -3 Croissance de régions.....</i>	<i>20</i>
<i>VI-4 Segmentation par approche contours</i>	<i>24</i>
<i>VI-4-1 La détection de contours</i>	<i>25</i>

VI-4-2 Fermeture des contours.....	26
VI-5 Les méthodes du split and merge.....	27
VI-6 Ligne de partage des eaux.....	29
VI-7 Méthodes SLIC.....	33
Conclusion.....	34
Chapitre II: La Saillance visuelle	
I- Introduction.....	37
II- Système visuel humain.....	37
III- Définition de la saillance.....	39
VI- Les facteurs de la saillance visuelle.....	40
VI-1 - Facteurs physiques.....	40
VI-1.1 La saillance liée aux contrastes de luminosité et de couleur.....	40
VI-1.2 La saillance liée à une singularité dans un ensemble d'objets.....	41
VI-1.3 La saillance liée à la cohésion et la structuration de la scène.....	42
VI-2 Facteurs cognitifs.....	44
VI-2.1 La saillance liée à l'intention.....	44
VI-2.2 La saillance liée à l'attention.....	44
VI-2.3 La saillance liée à la mémoire à court terme.....	45
VI-2.4 La saillance liée à la mémoire à long terme.....	46
VI-2.5 La saillance liée à l'affect et aux émotions.....	46
V- Méthods de saillance: Bottom-up et Top-down.....	47
IV- Modèles de carte de saillance	48
IV-1 hiérarchique.....	48
IV-2 Théorie de l'information.....	50
IV-3 Connexionniste.....	50
IV-4 Algorithmique.....	51
IV-5 Graphique.....	52
Conclusion.....	53

Chapitre III choix de la méthode & mise en œuvre

<i>I- Présentation de l'environnement de travail.....</i>	<i>55</i>
<i>I-1 Présentation de MATLAB.....</i>	<i>55</i>
<i>II- Présentation de la méthode.....</i>	<i>57</i>
<i>II-1. Introduction.....</i>	<i>57</i>
<i>II-2 Multi-Graphes basé sur Manifold Ranking.....</i>	<i>60</i>
<i>II-3. La Construction de Graphe.....</i>	<i>62</i>
<i>II-5 Détection de saillance en deux étapes.....</i>	<i>63</i>
<i>II-5-1 Classement avec des requêtes d'arrière plan.....</i>	<i>64</i>
<i>II-5-2 Classement avec des requêtes d'avant plan.....</i>	<i>65</i>
<i>II-6 Algorithme : la saillance basée sur la méthode multi-graphes.....</i>	<i>66</i>
<i>II-7 Exemple d'application de l'algorithme.....</i>	<i>67</i>
<i>II-8 Résultats expérimentaux.....</i>	<i>67</i>
<i>Conclusion.....</i>	<i>71</i>
<i>Conclusion générale.....</i>	<i>73</i>
<i>Bibliographie.....</i>	<i>75</i>

Liste des Figures :

Chapitre I:

<i>Fig. 1: Exemple d'image en niveaux de gris.....</i>	<i>06</i>
<i>Fig. 2: Exemple d'image binaire.....</i>	<i>07</i>
<i>Fig. 3: Exemple d'image couleur</i>	<i>07</i>
<i>Fig. 4: Illustration d'un pixel.....</i>	<i>08</i>
<i>Fig. 5: Notion de voisinage d'un pixel.....</i>	<i>09</i>
<i>Fig. 6: Exemple d'un histogramme.....</i>	<i>10</i>
<i>Fig. 7: Image en niveau de gris et son image binarisée.....</i>	<i>11</i>
<i>Fig. 8: Segmentation d'une image couleur.....</i>	<i>15</i>
<i>Fig. 9: Différentes régions d'une image.....</i>	<i>15</i>
<i>Fig. 10: Histogramme de la répartition des différentes zones couleurs.....</i>	<i>16</i>
<i>Fig.11 : Etapes à suivre pour la segmentation par histogramme.....</i>	<i>17</i>
<i>Fig. 12: Exemple de seuillage par minimisation de variance.....</i>	<i>18</i>
<i>Fig. 13: Deux schémas de regroupement pour des pixels a, b et c.....</i>	<i>21</i>
<i>Fig. 14: Image originale avant accroissement de région.....</i>	<i>22</i>
<i>Fig. 15: Accroissement de région1.....</i>	<i>23</i>
<i>Fig. 16: Accroissement de région2.....</i>	<i>23</i>
<i>Fig. 17: Détection de contours sur Lena.</i>	<i>24</i>
<i>Fig. 18: Illustration de contours à fermer sur Lena.....</i>	<i>24</i>
<i>Fig. 19: Gradient d'une image à 2 régions.</i>	<i>26</i>
<i>Fig. 20: Subdivision d'une feuille.....</i>	<i>27</i>
<i>Fig. 21: Subdivisions successives jusqu'à convergence.....</i>	<i>27</i>
<i>Fig. 22: Graphe de Région Adjacence.....</i>	<i>28</i>
<i>Fig. 23: Fusion de nœuds du RAG.....</i>	<i>28</i>
<i>Fig. 24: Résultat du Split and Merge sur une image donnée.....</i>	<i>29</i>
<i>Fig. 25: La ligne de partage des eaux comme étant la crête formant la limite entre deux bassins versants.....</i>	<i>29</i>
<i>Fig. 26: Illustration de la ligne de partage des eaux.....</i>	<i>30</i>
<i>Fig. 27: Illustration de la sur-segmentation.....</i>	<i>31</i>
<i>Fig. 28: Image lena.....</i>	<i>32</i>

<i>Fig. 29: Image cell.....</i>	<i>32</i>
<i>Fig. 30: Images segmentées en utilisant SLIC.....</i>	<i>33</i>
<i>Fig. 31: Présentation de l'algorithme SLIC.....</i>	<i>34</i>

Chapitre II

<i>Fig. 1: Les champs récepteurs centre-périphérie.....</i>	<i>38</i>
<i>Fig. 2: Structure du modèle d'attention visuelle.....</i>	<i>48</i>
<i>Fig. 3: Le Modèle connexionniste.....</i>	<i>51</i>

Chapitre III

<i>Fig. 1: lustration de la version de MATALB.....</i>	<i>55</i>
<i>Fig. 2: fenêtre principale de MATALB.....</i>	<i>55</i>
<i>Fig. 3: Editeur de texte MATALB.....</i>	<i>56</i>
<i>Fig.4: Génération des quatre cartes de saillance.....</i>	<i>58</i>
<i>Fig. 5: Notre modèle de graphe.....</i>	<i>60</i>
<i>Fig. 6: Illustration des différentes étapes de notre algorithme.....</i>	<i>67</i>
<i>Fig. 7: Quelques résultats de notre travail.....</i>	<i>69</i>
<i>Fig. 8 : courbes précision rappel & la précision, le rappel et F-mesure.....</i>	<i>71</i>

INTRODUCTION GENERAL :

La vision est incontestablement le sens le plus développé chez l'homme : Les humains ont une capacité remarquable d'analyser et d'interpréter en temps réels des scènes complexes. Par ailleurs, la vision par ordinateur est devenue ces dernières années un domaine en plein essor de part ses multiples applications en médecine, robotique, surveillance. . . . Ainsi, l'étude du système visuel humain ou animal dans le but d'en comprendre et copier les mécanismes pour la vision par ordinateur est particulièrement intéressante. Diverses analyses des systèmes visuels animaux ont été menées ces trente dernières années. Elles montrent que le regard est en premier lieu attiré par certaines caractéristiques d'une scène. C'est ce que l'on appelle l'attention visuelle, qui est la capacité d'un système à sélectionner rapidement les informations les plus pertinentes du champ de vision. Des recherches ont été conduites pour mettre en place un système informatique d'attention visuelle proche de celui utilisé par les humains. Une partie des modèles d'attention visuelle proposés jusqu'à présent est basée sur les connaissances physiologiques des primates. Chez eux, les zones attentionnelles sont sélectionnées en combinant un balayage de la scène indépendamment du but recherché, faisant juste ressortir les points saillants (processus bottom-up) avec un balayage dépendant du but recherché et nécessitant des connaissances préalables (processus top-down). Ces systèmes d'attention visuelle ont pour rôle principal l'accélération et l'amélioration des modèles de vision artificielle. Les applications possibles sont diverses : compression d'images, détection d'objets ou même amélioration des panneaux publicitaires....

Le but de notre travail est de construire un système de direction d'attention visuelle, basé sur le mécanisme qui pilote les fixations visuelles humaines, en ne considérant que le processus bottom-up. Ce système doit être capable d'analyser n'importe quelle image présentée et de retourner les coordonnées des points les plus saillants. Le travail consiste à reprendre un modèle développé par Chuan Yang et Xiang Ruan qui est basé sur le graphe Manifold Ranking, à l'analyser et à le modifier, en particulier en détectant les points saillants à différentes échelles. Ce mémoire présente l'ensemble du travail et des résultats obtenus au cours de cette recherche. Il

Introduction Général

s'articule autour du plan suivant. Un premier chapitre qui se portera sur les généralités sur les traitements et la segmentation d'images ensuite un deuxième chapitre qui dont on parlera sur la saillance visuelle et ses différentes approches et puis dans le troisième chapitre nous allons présenter notre application et ses résultats obtenus tout en les comparons avec d'autres méthodes d'état de l'art. Ce rapport se termine par la présentation et l'évaluation du système obtenu à partir de ces modifications.

Chapitre I

***GENERALITES SUR LE
TRAITEMENT***

&

***SEGMENTATION
D'IMAGE***

Partie I :
Généralités sur le traitement
d'images

I. INTRODUCTION :

L'image constitue l'un des outils les plus importants qu'utilise l'être humain pour communiquer avec autrui. C'est un moyen de communication universel dont la richesse du contenu permet aux individus de tout âge et de toute culture de se comprendre.

Chacun peut alors analyser l'image à sa manière. Pour en dégager une impression et en extraire des informations précises. De ce fait, le traitement d'images est l'ensemble des méthodes et techniques permettant d'améliorer l'aspect visuel d'une image.

Ainsi, dans ce chapitre, nous allons présenter les notions générales sur l'image et les concepts qui s'y rattachent à ce domaine.

II- NOTIONS D'IMAGE :

L'image est une représentation d'objets ou de personnes par la peinture, la sculpture, le dessin, la photographie, le film ou divers procédés électroniques de visualisation [J.P Coquerez.95]. C'est aussi une projection à deux dimensions d'une scène à trois dimensions. Elle est définie comme une matrice où la valeur de chaque élément (pixel) représente l'intensité discrète de la lumière en ce point. Chaque cellule élémentaire de l'image (pixel) a comme caractéristique un niveau de gris ou une valeur colorimétrique prélevée à l'emplacement correspondant dans l'image réelle [M.Ouali. K.Mohamedi.09].

Il existe trois types d'image : image en niveau de gris, image binaire et image en couleur.

II.1. Image en niveau de gris :

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires [F.Boudjemai.09] (*Fig. 1*).

La représentation des images en niveaux de gris, fait en attribuant à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise entre 0 et 255. Chaque pixel est donc représenté par un octet.



Fig.1 Exemple d'image en niveaux de gris

La conversion d'une image couleur en une image en niveau de gris se fait plus exactement en négligeant les composantes de saturation et de teinte d'une image et en ne tenant compte que de la composante de luminance. Ceci peut être réaliser par une somme pondérée des composantes R,G,B suite à l'application de la fonction suivante :

$$I(i,j) = \alpha * R + \beta * G + \sigma * B \quad (1)$$

Avec :

i,j : l'emplacement spatiale du pixel.

α, β, σ : ce sont des points à définir préalablement.

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

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

$G(i,j)$:intensité de la couleur verte du pixel(i,j)

$B(i,j)$:intensité de la couleur bleue du pixel(i,j)

II.2. Image binaire :

Une image binaire est une matrice dont les niveaux de gris sont réduits aux deux éléments 0 et 1 Le niveau 0 représente le noir absolu et le niveau 1 représente le blanc [F.Boudjemai.09]. (**Fig.2**). Dans ce cas un seul bit est utilisé pour coder le niveau de gris.



Fig. 2 Exemple d'image binaire

II.3. Image couleur :

Les applications multimédias utilisent le plus souvent des images en couleurs (Figure 3). La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités.

En effet, il faut tout d'abord choisir un modèle de représentation (par exemple en RGB,). On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange de rouge, vert et bleu (modèle R.G.B.):



Fig 3 : Exemple d'image couleur

II.4. Caractéristiques de l'image :

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

a. Pixel :[J.Dubois]

Un pixel est l'unité de base d'une image numérique. Il constitue l'unité minimale adressable par le contrôleur vidéo. A chaque pixel est associé une couleur, elle même décomposée en 3 composantes primaires qui sont le Rouge, le Vert et le Bleu (RGB), la (**Fig.4**) montre un exemple de pixel :

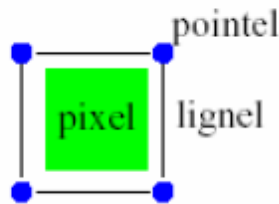


Fig.4 : illustration d'un pixel

Ici, la couleur associée est le vert et on voit bien la forme rectangulaire voire carrée de cet élément de base d'une image numérique.

b. Notion de voisinage d'un pixel : [J.P.Coquerez.95]

Dans une image, un pixel $p(x,y)$ possède quatre voisins (horizontaux et verticaux) qui soumit ce qu'on appelle le 4-voisinage de $p(x,y)$. Si on considère un pixel comme un carré élémentaire, les pixels présentant un cote commun avec $p(x,y)$ sont appelés le 4-voisins de $p(x,y)$. Le pixel $p(x,y)$ possède aussi quatre voisins diagonaux, ce sont les pixels qui ont un sommet commun avec $p(x,y)$. L'ensemble des huit voisins du pixel (x,y) représente ce qu'on appelle les 8-voisins (8- voisinage). Ces huit voisins forment la fenêtre 3x3 du pixel $p(x,y)$.

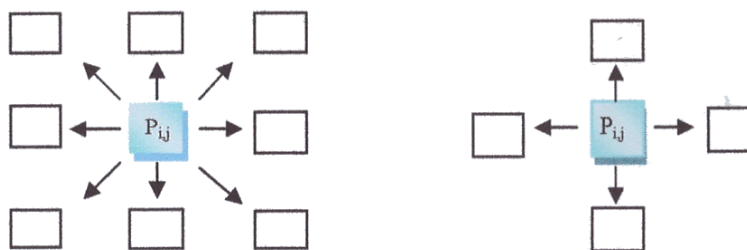


Fig .5 Notion de voisinage d'un pixel

Gauche : 8-voisinage

Droite : 4-voisinage

c. Dimension

La dimension est la taille de l'image, qui se présente sous forme de matrice dont les éléments sont des valeurs numériques représentatives des intensités lumineuses (pixels). Le nombre de lignes de cette matrice multiplié par le nombre de colonnes nous donne le nombre total de pixels dans une image

d. Résolution

La résolution est exprimée en nombre de pixels par unité de mesure (pouce ou centimètre). Le mot résolution est utilisé pour désigner le nombre total de pixels affichables horizontalement ou verticalement sur un moniteur; plus le nombre de pixel est grand, meilleure est la résolution.

e. Bruit

Un bruit (parasite) dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins [C.Kaddour]

f. La luminance

C'est le degré de luminosité des points de l'image [C.Kaddour]. 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 [M.Ouali. K.Mohamed.09].

i. Le contraste

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

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

III. HISTOGRAMME D'UNE IMAGE :

L'histogramme d'une image est représenté par un graphique, l'axe horizontal (abscisse) représente l'ensemble des valeurs de luminosité potentielles de l'image : c'est la quantité de lumière (ou l'intensité lumineuse) sur 256 niveaux, allant de 0 (le noir) à

Chapitre I : Généralités sur le traitement & segmentation d'image

255 (le blanc). L'axe vertical (ordonnée) représente la proportion de pixels dans l'image pour chacun des niveaux de luminosité (*Fig.6*).

Les tons foncés sont représentés à gauche de l'histogramme. Plus on va vers la droite, plus on augmente la luminosité. De même, lorsque l'histogramme monte, le nombre de pixels augmente. Lorsqu'il descend, le nombre de pixels diminue. Si un grand nombre de pixels se situent sur la partie gauche de l'histogramme, cela signifie qu'une majorité de pixels se trouvent dans les basses lumières.

Il peut être utilisé pour améliorer la qualité d'une image (Rehaussement d'image) en introduisant quelques modifications, pour pouvoir extraire les informations utiles de celle-ci [C.Kaddour].

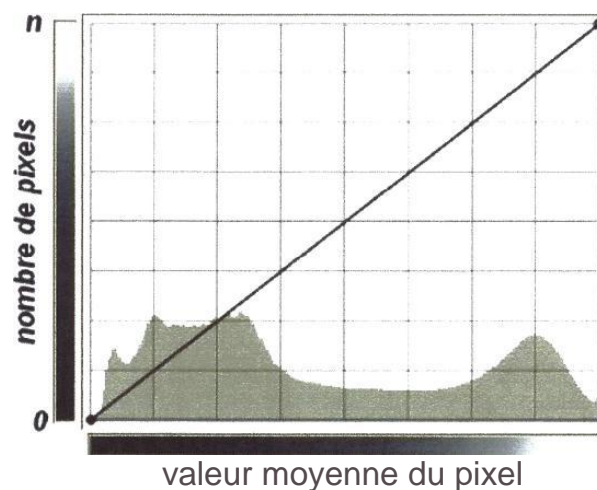


Fig.6 : Exemple d'un histogramme

Notons que, pour les images en couleurs, plusieurs histogrammes sont nécessaires, pour une image RGB (Rouge, Green, Bleu) par exemple les histogrammes représentant respectivement la distribution des valeurs des composantes rouge vert et bleu sont nécessaires [M.Ouali. K.Mohamedi.09].

Binarisation :

La binarisation consiste à obtenir à partir d'une image codée sur plusieurs niveaux de gris (généralement 256), une image codée sur deux niveaux de gris (noir et blanc). Elle a comme but de diminuer la quantité d'informations présentes dans l'image, et de ne garder que les informations pertinentes [H.Cettach.05]. La binarisation permet de mettre en évidence des formes ou des objets dans une image. Le principe de la binarisation consiste à déterminer un seuil S appelé seuil de binarisation,

Chapitre I : Généralités sur le traitement & segmentation d'image

puis modifier les valeurs des pixels de l'image comme suit

Soit I une image et $I(x, y)$ un point de I de coordonnées (x, y)

$$I(x,y) = \begin{cases} 0 & \text{Si } I(x,y) < S \\ 255 & \text{Sinon} \end{cases} \quad \begin{matrix} (\text{noir}) \\ (\text{blanc}) \end{matrix} \quad (3)$$



a : Image en niveau du gris

b : Image binarisée

Fig.7 Image en niveau de gris et son image binarisée

IV. LES FILTRES :

Afin d'améliorer la qualité visuelle de l'image, on doit éliminer les effets des bruits (parasites) en lui faisant subir un traitement appelé filtrage [M.Ouali. K.Mohamedi.09].

On distingue généralement les types de filtres suivants :

- *Filtres passe-bas (lissage)* : Ce type de filtrage consiste à atténuer les composantes de l'image ayant une séquence haute (pixels fonces), il est utilisé généralement pour atténuer le bruit de l'image. C'est la raison pour laquelle on parle habituellement de lissage.
- *Filtres passe-haut (accentuation)* : à l'inverse des passe-bas, ils atténuent les composantes de basse séquence de l'image et permettent, notamment d'accentuer des contours et leur extraction.

Chapitre I : Généralités sur le traitement & segmentation d'image

CONCLUSION :

Dans ce chapitre, nous avons fait un aperçu général sur l'image et sur les différents traitements qui lui sont appliqués afin d'exploiter, de façon optimale, les informations qu'elle contient.

Dans ce qui suit, nous allons nous focaliser sur l'étape de segmentation, qui est une étape incontournable dans tout processus de traitement d'images. Elle permet de distinguer les différents composants d'une image (zones homogènes, formes, contours séparent différentes zones) pour une meilleure interprétation des informations pertinentes qu'elle renferme et ces différentes approches.

Partie II :
La segmentation d'images

I- INTRODUCTION :

La segmentation d'images est un problème fondamental mais encore difficile en vision par ordinateur et traitement d'image. C'est un processus essentiel pour de nombreuses applications telles que la reconnaissance d'objet, la poursuite de cible, la recherche d'image par le contenu et le traitement d'images médicales, ...etc. D'une manière générale, l'objectif de la segmentation de l'image est de partitionner une image en un certain nombre de pièces qui ont des caractéristiques cohérentes (couleur, texture, etc.) et dans le but de regrouper l'ensemble de pièces significatives pour des raisons de commodité de perception [Forsyth.Ponce.02]. Dans de nombreuses applications pratiques, comme il y a un grand nombre d'images qui sont manipulées et traitées, les interactions humaines appliquées dans le processus de segmentation doivent être le moins possibles. Cela rend les techniques de segmentation automatiques de l'image plus attrayantes. En outre, le succès de nombreuses techniques de segmentation de haut niveau (par exemple, la segmentation fondée sur la classe d'objet [L.Ladicky et al.09] et [F.Lecuberry et al.10]) exige également des techniques de segmentation automatiques sophistiquées. C'est une des étapes critiques de l'analyse d'images qui conditionne la qualité des mesures effectuées ultérieurement. Elle permet d'isoler dans l'image les objets sur lesquels doit porter l'analyse, de séparer, dans une image, les régions d'intérêt du fond.

II- DEFINITIONS :

II- 1. Définition de la segmentation [J.Dubois] :

La segmentation consiste à partitionner l'image étudiée en régions disjointes avec des couleurs et textures homogènes. La segmentation peut être basée sur différents critères comme donc la couleur, la texture ou les niveaux de gris d'une image. Voici un exemple de segmentation d'images permettant de montrer ce que l'on entend par régions homogènes par rapport au critère de couleurs [Corquez.Philipp.95] :

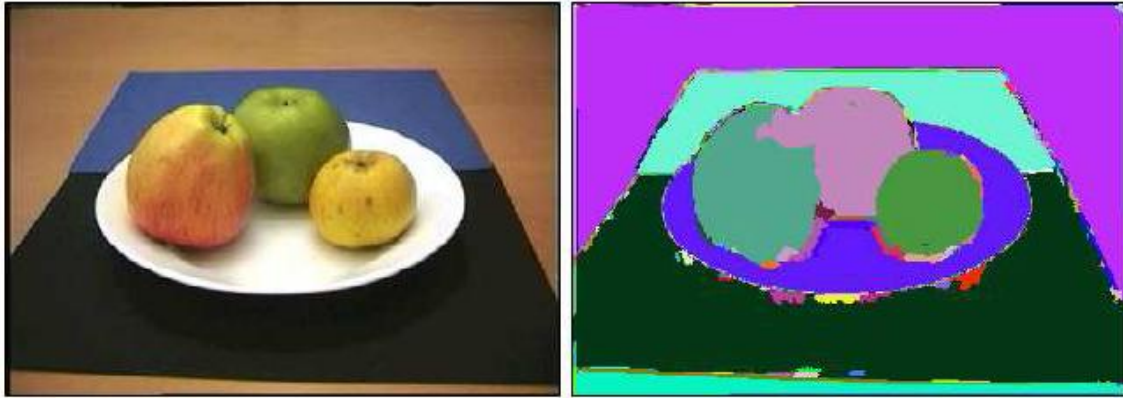


Fig.8 : Segmentation d'une image couleur

Sur cette image, on voit bien que chaque objet de l'image se voit attribué une couleur, d'où la séparation en régions dites homogènes. Cependant on voit qu'il peut y avoir des défauts de reconnaissance et donc y avoir des confusions entre les régions comme c'est le cas ici entre la pomme du milieu et celle se situant à sa gauche car elles ont une zone en violet commune.

II-2. Définition d'une région [J.Dubois]

La notion de région dans le processus de segmentation d'images, est de regrouper des zones possédant les mêmes caractéristiques. C'est-à-dire que si plusieurs pixels adjacents s'avèrent être de couleur identique alors la zone qu'ils forment est une région. Ci-après, une illustration montrant deux régions de pixels différentes :

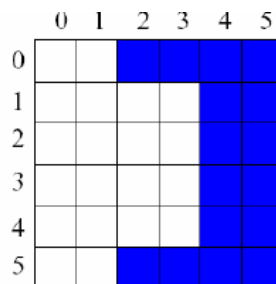


Fig .9 : Différentes régions d'une image

III. APPROCHES DE LA SEGMENTATION :

Il existe plusieurs types d'approches pour réaliser une segmentation. En effet, ce nombre d'approches s'élève au nombre de trois, à savoir :

- Les approches globales.
- Les approches locales.

- Les approches hybrides.

III- 1. Approche globale de la segmentation :

Le principe de cette approche c'est de détecter les pixels rare puis les comparer et faire des statistiques pour les regrouper selon un critère donné (couleurs, textures.....). Autrement dit, il s'agit de trouver dans l'image les régions qui correspondent au concept dont on dispose [M.Sonka et all.08].

III-2. Approche locale de la segmentation :

Cherche à définir des segments dans une image en y faisant ressortir des régions homogènes, c'est-à-dire des régions dont les pixels sont semblables selon certains critères. Dans ce cas, la segmentation est complètement guidée par les données [M.Sonka et all.08].

III-3. Approche hybride de la segmentation (mixte) :

C'est une approche qui fusionne le principe des deux approches citées ci-dessus.

VI- LES METHODES DE SEGMENTATION :

VI -1 Les histogrammes : [J.Dubois]

C'est une méthode qui se base sur l'approche globale, le principe de ce type de segmentation est de réaliser une segmentation sur un espace dérivé de l'image elle-même et d'en extraire par exemple une couleur. C'est une approche dite globale car l'histogramme calculé est basé sur toute l'image et chaque couleur correspond à un pic à l'intérieur de ce dernier comme l'illustration ci-dessous peut le mettre en évidence :

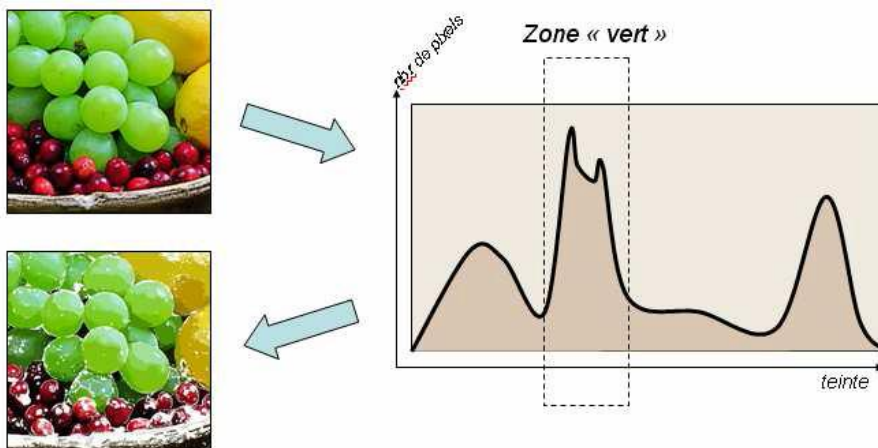


Fig.10 : Histogramme de la répartition des différentes zones couleurs

Chapitre I : Généralités sur le traitement & segmentation d'image

Ici, on voit bien la zone de la couleur verte sur l'histogramme qui correspond, en effet, a un pic.

Ensuite, chaque pixel de l'image est décrit selon un canal d'un espace de couleur. Par exemple, pour l'espace RGB le pixel est projeté dans le canal R, puis le G et enfin le B. Cela est exactement pareil si on choisit l'espace YUV par exemple.

Ainsi, on a les différentes étapes suivantes pour parvenir à isoler la couleur verte de cette image :

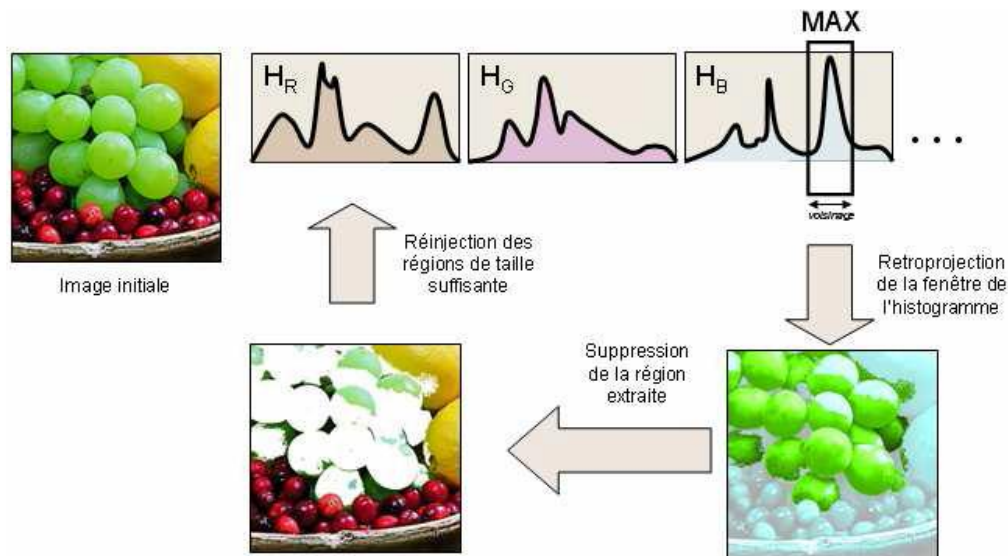


Fig 11 : Etapes à suivre pour la segmentation par histogramme

Ici, on voit comment isoler la couleur verte avant de réinjecter dans l'image les autres couleurs à extraire.

Minimisation de variance : La répartition des pixels en N classes est un problème classique de classification. Le choix des seuils S_i permet de détecter m classes auxquels on peut associer taille (t_i), moyenne (m_i) et variance V_i par :

$$t_i = \text{Sum}_{D_j} \{H(j)\} \quad (4)$$

$$m_i = \text{Sum}_{D_j} \{j \cdot H(j) / t_i\} \quad (5)$$

$$v_i = \text{Sum}_{D_j} \{(j - m_i)^2 \cdot H(j) / t_i\} \quad (6)$$

Où H est l'histogramme normalisé (son intégrale est ramenée à l'unité) et $D_j = [s_{i-1}, s_i[$ est la gamme de niveaux de gris correspondant à la classe C_i (par hypothèse, $s_0 = 0$) .

A partir de ces indicateurs statistiques, on peut construire la variance intercalasse totale W par :

$$W = \sum_i \{t_i \cdot V_i\} \quad (7)$$

Le meilleur seuillage dans cette approche correspond à une minimisation de la variance intra classe. Cette technique est difficilement applicable lorsque le nombre de classes est élevé. En effet, il faut tester exhaustivement tous les $(N-1)$ -uples (s_1, \dots, s_{N-1}) possibles. De plus, il faut que chaque classe ait une taille significative en nombre de niveaux de gris pour que les indicateurs statistiques aient un sens.

Plus récemment, [N.Otsu.79] a proposé de réaliser une maximisation de la variance inter-classe qui, dans le cas de la binarisation, s'exprime par

$$B = t_0 \cdot t_1 (m_0 - m_1)^2 \quad (8)$$

Ce qui est rigoureusement équivalent puisque on a la relation :

$$W + B = cste \quad (9)$$

Cependant, la méthode d'Otsu est plus intéressante d'un point de vue calculatoire car elle ne nécessite pas de calcul de variances.



Fig 12. Exemple de seuillage par minimisation de variance

Sur cet exemple, un seuil a été déterminé par l'algorithme optimal de [R.fisher.95] Le seuil trouvé est 127. Compte tenu du fait que l'on n'utilise qu'un seul seuil, toutes les composantes de l'image ne peuvent être séparées.

Seuillage entropique : Le seuillage entropique est une technique dérivée de la théorie de l'information. Les seuils sont déterminés de manière à maximiser l'entropie E résultant du découpage de l'histogramme H en plusieurs classes. En effet, l'entropie

Chapitre I : Généralités sur le traitement & segmentation d'image

mesure la quantité d'information portée par un groupe. Pour un nombre de seuils fixe, on cherche à ce que les classes résultantes portent le maximum d'information.

L'entropie totale est calculée de la manière suivante:

$$E = \text{Sum}_i E(C_i) \quad (10)$$

Où C_i désigne la classe N_{oi} ,

$$E(C_i) = -\text{Sum}_{D_j} \{p_j \cdot \log_2(p_j)\} \quad (11)$$

où D_j est l'ensemble des niveaux de gris j associés à la classe C_i et p_j la probabilité a postériori du niveau de gris j , estimée par

$$p_j = H(j) / \text{taille de l'image} \quad (12)$$

La notion d'entropie n'est pas liée à une caractéristique visuelle. C'est pourquoi l'image résultat paraît le plus souvent de moins bonne qualité (du point de vue de l'œil humain).

VI -2- Méthodes k-means: [J.Mcqueen.67]

C'est une méthode qui se base sur le principe de l'approche Top-Down, elle représente un outil de classification classique qui permet de répartir un ensemble de données en K classes homogènes. La plupart des images vérifient localement des propriétés d'homogénéité, notamment en termes d'intensité lumineuse. L'algorithme des k -means permet donc d'apporter une solution à la segmentation d'images. L'algorithme k -means est l'algorithme de clustering le plus connu et le plus utilisé, du fait de sa simplicité de mise en œuvre. Il partitionne les données d'une image en K clusters. Contrairement à d'autres méthodes dites hiérarchiques, qui créent une structure en « arbre de clusters » pour décrire les groupements, k -means ne crée qu'un seul niveau de clusters. L'algorithme renvoie une partition des données, dans laquelle les objets à l'intérieur de chaque cluster sont aussi proches que possible les uns des autres et aussi loin que possible des objets des autres clusters. Chaque cluster de la partition est défini par ses objets et son centroïde.

Le k -means est un algorithme itératif qui minimise la somme des distances entre chaque objet et le centroïde de son cluster. La position initiale des centroïdes conditionne le résultat final, de sorte que les centroïdes doivent être initialement placés le plus loin possible les uns des autres de façon à optimiser l'algorithme. K -means

Chapitre I : Généralités sur le traitement & segmentation d'image

change les objets de cluster jusqu'à ce que la somme ne puisse plus diminuer. Le résultat est un ensemble de clusters compacts et clairement séparés, sous réserve qu'on ait choisi la bonne valeur K du nombre de clusters.

Les principales étapes de l'algorithme *k-means* sont :

1. Choix aléatoire de la position initiale des K clusters.
2. (Ré-) Affecter les objets à un cluster suivant un critère de minimisation des distances (généralement selon une mesure de distance euclidienne).
3. Une fois tous les objets placés, recalculer les K centroïdes.
4. Répéter les étapes 2 et 3 jusqu'à ce que plus aucune réaffectation ne soit faite.

VI -3 Croissance de régions : [J.Dubois]

Les méthodes d'accroissement de région sont les méthodes de segmentation les plus simples. Le principe est basé sur une approche bottom-up : l'algorithme part de petits éléments de l'image qu'il va tenter de regrouper en éléments plus importants.

Nous présentons ici la version de base de l'algorithme d'accroissement de région qui fonctionne en agrégeant des pixels.

Principe de fonctionnement :

Supposons une région de couleur homogène R . Initialement, $R = 1 \text{ pixel}$.

On va étendre la région R en incluant les pixels situés sur la frontière et dont la couleur est proche de celle de R (la variation de couleur est inférieure à un seuil ; caractéristique de ce type d'algorithmes). En répétant cette procédure jusqu'à ce qu'il n'y ait plus de pixels de couleur assez proche sur la frontière, on obtient une région de couleur homogène maximale autour du pixel de départ. La région initiale 'gonfle'. en absorbant des pixels de la frontière, jusqu'à stabilité par rapport à la propriété d'homogénéité.

Cette méthode présente deux limitations sévères qui n'en font pas une méthode très efficace :

1. Les régions obtenues dépendent fortement des pixels d'amorçage choisis et de l'ordre dans lequel les pixels de la frontière sont examinés.
2. Le résultat final est très sensible à la valeur du seuil δ :

Chapitre I : Généralités sur le traitement & segmentation d'image

Pour illustrer le premier problème, considérons trois pixels adjacents a , b et c dont les intensités respectives sont 8, 10 et 11 (par exemple, l'intensité en niveaux de gris). Le seuil est 2. La région initiale est constituée du pixel b . Deux schémas de regroupement pour les points frontière a et c sont possibles (**Fig.13**)

Le pixel central b est l'amorce. Compte tenu du seuil $\delta = 2$, a et c sur la frontière devraient être ajoutés à l'amorce b . Si l'on commence par tenter d'agréger a , le résultat du regroupement, noté $[ab]$, a pour intensité moyenne 9 et c s'y ajoute ensuite puisque $|11 - 9| \leq \delta$. On a donc regroupé a , b et c . Si maintenant l'algorithme commence par examiner le point frontière c , le groupement de b et c donne $[bc]$ dont l'intensité est 10,5. Le point frontière a d'intensité 8 est trop éloigné et il est considéré comme appartenant à une autre région. On obtient donc deux groupements au lieu d'un.

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

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

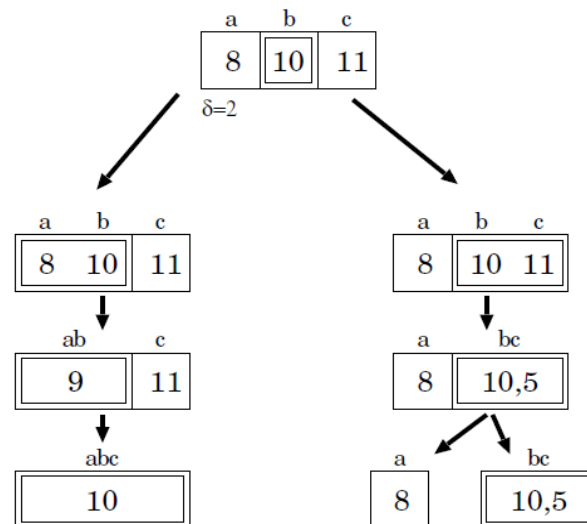


Fig. 13. Deux schémas de regroupement pour des pixels a , b et c

Cet algorithme fait parti de la classe d'algorithmes de segmentation dit « locaux ». L'opération élémentaire consiste à manipuler des pixels adjacents et l'algorithme n'a aucune vision globale du résultat qu'il obtient. Par exemple, il est

Chapitre I : Généralités sur le traitement & segmentation d'image

incapable de détecter que la région qu'il vient de construire est inhomogène, ce qui est souvent le cas.

Exemples

Voici un exemple dont l'image test a été empruntée au site internet de L'Image Recognition Laboratory de l'université Koblenz-Landau (*fig.14*).



Fig. 14 Image originale avant accroissement de région

Un algorithme d'accroissement de région a été appliqué avec un seuil de 6 puis de 7 sur le maximum de variation de RGB ($\max(\Delta R; \Delta G; \Delta B) < \delta$).

Les images segmentées sont présentées en fausses couleurs c'est à dire que chaque région est coloriée d'une manière aléatoire de sorte qu'on parvienne à la distinguer nettement des régions adjacentes (*fig15*).

A gauche, l'image avec un seuil de 6. On constate que les fleurs sont correctement extraites mais la base en bas se confond avec le fond. Ce problème est caractéristique des algorithmes de segmentation dit « locaux » qui n'ont pas de vue d'ensemble ou de description de plus haut niveau que le pixel. Dans le cas de l'accroissement de région, entre deux points de deux régions connexes, il existe presque toujours une chaîne de pixels connexes pour laquelle l'écart de couleur entre deux pixels adjacents est inférieur au seuil. Cela se traduit par des effets de « ponts » entre zones, particulièrement sensibles dans les zones de dégradé. La conséquence est que les régions ont tendance à se mélanger au delà d'un certain seuil, et ceci même si les régions en question sont très différentes du point de vue de la couleur. Par exemple,

Chapitre I : Généralités sur le traitement & segmentation d'image

il est facile de créer un pont de pixels entre un carré rouge et un carré bleu en utilisant un dégradé continu de rouge vers bleu.

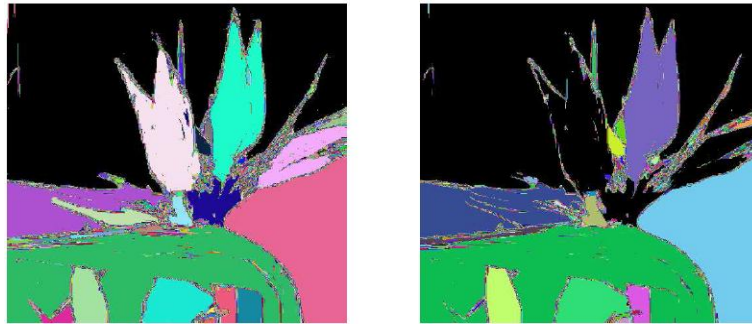


Fig.15 *Accroissement de région1 : A gauche, seuil=6. A droite, seuil=7*

A droite, la même image est segmentée avec un seuil de 7. On constate que la fleur du milieu est confondue avec le fond (il existe visiblement un pont de pixels pour le seuil 7 entre cette région et le fond). Ceci met en lumière l'un des principaux problèmes liés à l'accroissement de région : la sensibilité au seuil. En pratique, il est presque toujours possible de trouver un seuil qui segmente bien une image donnée. Le problème vient que si l'on augmente d'une unité ce seuil, l'image est mal segmentée (fusion de régions, problème des « ponts »), si on diminue ce seuil d'une unité, l'image est sur-segmentée. Il n'existe pas de seuil qui fonctionne correctement pour toute une classe d'images, ce qui rend l'automatisation des algorithmes très difficile. A titre d'exemple, la (**fig.16**) montre la segmentation de l'image précédente avec un seuil de 3 (à gauche) et de 4 (à droite). Le résultat est correct pour le seuil de 4 et devient beaucoup trop morcelé pour le seuil de 3. Le seuil optimal de cette image est donc 4. Pour une autre image, la bonne valeur de seuil sera en général différente et il faut recommencer ces estimations, ce qui rend l'algorithme peu utilisable dans un environnement bruité et complexe.

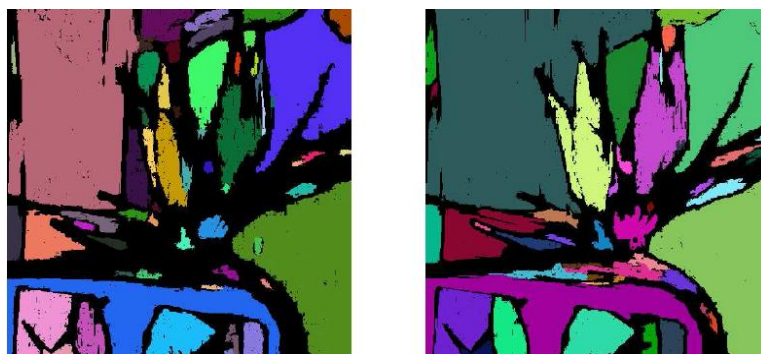


Fig.16 *Accroissement de région2 : A gauche, seuil=3 A droite, seuil=4*

VI-4 Segmentation par approche contours : [J.Dubois]

La segmentation par approche contours [Ludovic.Macaire] et [Alexandre.Hamon] s'intéresse aux contours de l'objet dans l'image. La plupart des algorithmes qui lui sont associés sont locaux, c'est-à-dire qu'ils fonctionnent au niveau du pixel.

Des filtres détecteurs de contours sont appliqués à l'image et donnent généralement un résultat difficile à exploiter sauf si les images sont très contrastées.

Les contours extraits sont la plupart du temps morcelés et peu précis, il faut alors utiliser des techniques de reconstruction de contours par interpolation ou connaître a priori la forme de l'objet recherché.

Formellement, ce type d'algorithme est proche des méthodes d'accroissement de régions fonctionnant au niveau du pixel. Ces techniques purement locales sont en général trop limitées pour traiter des images bruitées et complexes.

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

- caractériser la frontière entre les régions :



Fig.17 : Détection de contours sur Lena

- fermer les contours :



Fig.18 : Illustration de contours à fermer sur Lena

VI-4-1. La détection de contours :

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

Les contours dans une image proviennent des :

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

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

- Approche gradient : détermination des extrema locaux dans la direction du gradient.
- Approche laplacien : détermination des passages par zéro du laplacien.

Le gradient d'une image :

Soit I : l'image à traité
 x, y : coordonnées d'image
 h : filtre de l'image

Le gradient d'une image est le vecteur $\nabla I(x, y)$ défini par :

$$\nabla I(x, y) = \left(\frac{\partial I(x, y)}{\partial x}, \frac{\partial I(x, y)}{\partial y} \right)^t \quad (14)$$

Il est donc caractérisé par un module m et une direction \emptyset dans l'image :

$$m = \sqrt{\left(\frac{\partial I(x, y)}{\partial x} \right)^2 + \left(\frac{\partial I(x, y)}{\partial y} \right)^2}, \quad (15)$$

$$\emptyset = \arctan \left(\frac{\partial I(x, y)}{\partial y} / \frac{\partial I(x, y)}{\partial x} \right). \quad (16)$$

La direction du gradient maximise la dérivée directionnelle.

La dérivée de $I(x, y)$ dans une direction donnée d s'écrit :

$$\nabla I(x, y) \cdot d. \quad (17)$$

Le gradient d'une image filtrée a pour définition :

$$\nabla \hat{I}(x, y) = \nabla (I(x, y) * h(x, y)) = \nabla I(x, y) * h(x, y) = I(x, y) * \nabla h(x, y). \quad (18)$$

Chapitre I : Généralités sur le traitement & segmentation d'image

Voici une illustration montrant le gradient d'une image découpé en deux zones distinctes

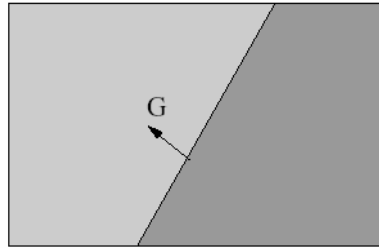


Fig.19 : Gradient d'une image à 2 régions

Le gradient n'est autre qu'une dérivée vectorielle de l'image et permettent de détecter les contours du fait que les contours correspondent à des discontinuités d'ordre 0 de la fonction d'intensité.

Cependant, le calcul de la dérivée nécessite un pré filtrage de l'image étudiée.

Le Laplacien d'une image :

Le laplacien d'une image d'intensité $I(x,y)$ est défini par :

$$\nabla^2 = \frac{\partial^2 I(x,y)}{\partial x^2} + \frac{\partial^2 I(x,y)}{\partial y^2} \quad (18)$$

- Invariant aux rotations de l'image.
- Le laplacien est souvent utilisé en amélioration d'images pour accentuer l'effet de contour :

$$I'(x,y) = I(x,y) - c \nabla^2 I(x,y) \quad (19)$$

- Sensibilité au bruit accrue par rapport au gradient.
- La Laplacien d'une image filtrée :

$$\Delta I'(x,y) = \Delta I(x,y) * h(x,y) = I(x,y) * \Delta h(x,y) \quad (20)$$

VI-4-2 Fermeture des contours :

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

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

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

Chapitre I : Généralités sur le traitement & segmentation d'image

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

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

VI-5 Les méthodes du split and merge : [J.Dubois]

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

Tout d'abord, voyons la première phase de cette méthode, le split :

Pour cette phase, l'image est stockée dans un arbre avec Initialement :

arbre racine = image complète

Récurssivement, chaque feuille F est subdivisée en quatre si elle n'est pas assez homogène, et les quatre sous images sont ajoutées en tant que feuilles de F, comme on peut le voir ci-dessous :

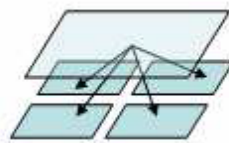


Fig.20 : Subdivision d'une feuille

L'algorithme poursuit tant qu'il reste des feuilles non homogènes à diviser jusqu'à arriver à la fin et d'avoir une structure du type :

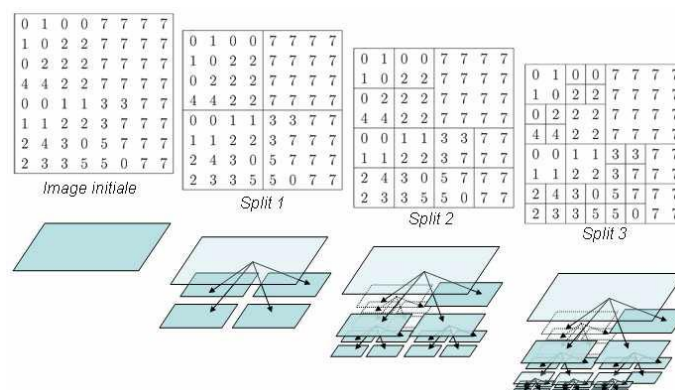


Fig.21 : Subdivisions successives jusqu'à convergence

Chapitre I : Généralités sur le traitement & segmentation d'image

L'étape suivante consiste à créer un graphe d'adjacence traduisant les découpages réalisés auparavant et qui servira par la suite dans la phase du merge :

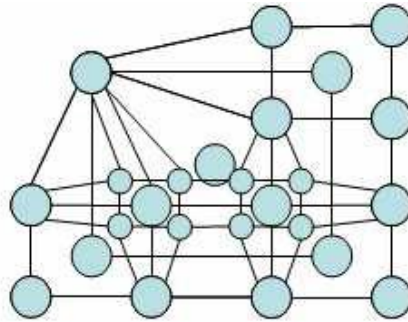


Fig.22: Graphe de Région Adjacence

Une fois le graphe d'adjacence déterminé, on peut passer à la phase du **merge** :

- Chaque nœud du Région Adjacence Graph (graphe d'adjacence) est examiné.
- Si un des voisins de ce nœud est à une distance inférieure à un seuil de regroupement, les deux nœuds fusionnent dans le RAG.
- Lorsque plus aucun nœud ne peut fusionner avec l'un de ses voisins, alors cela entraîne une condition d'arrêt pour l'algorithme comme l'illustre le schéma ci-dessous:

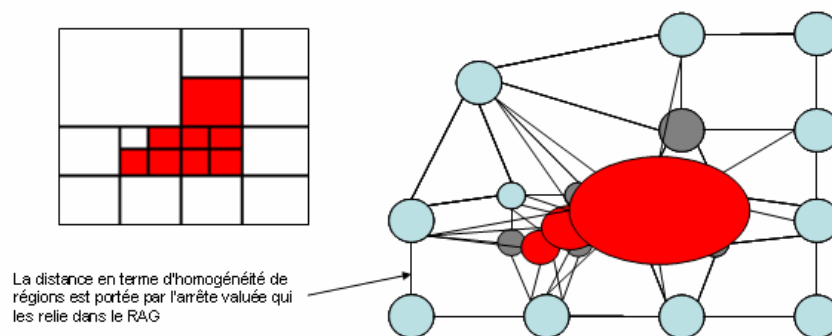


Fig.23 : Fusion de nœuds du RAG

Ici, on voit bien les étapes de fusion avec des voisins (représenté en rouge). Cette méthode comporte certains avantages comme celui de s'affranchir du problème du gradient, comme nous pourrions le voir dans l'approche locale de la segmentation, mais a pour défaut principal d'être complexe à mettre en œuvre et de provoquer des découpages trop carres qui sont dus à la nature des graphes.

Voici une illustration mettant en évidence ce dernier problème :



Fig 24 : Résultat du Split and Merge sur une image donnée

En effet, on peut constater que sur l'image ayant « subi » le split and merge, une zone carrée est apparue sur l'image près des panneaux de signalisation alors qu'elle n'apparaissait pas sur l'image originale.

VI-6 Ligne de partage des eaux¹ :

C'est des méthodes qui se basent sur le principe de l'approche hybride, en morphologie mathématique, l'algorithme de la Ligne de Partage des Eaux (LPE), proposé par Digabel et Lantuéjoul en 1972, est un algorithme de segmentation d'images, ce qui revient à décomposer une image en régions homogènes.

La ligne de partage des eaux utilise la description des images en termes géographiques. Une image peut en effet être perçue comme un relief si l'on associe le niveau de gris de chaque point à une altitude. Il est alors possible de définir la ligne de partage des eaux comme étant la crête formant la limite entre deux bassins versants.

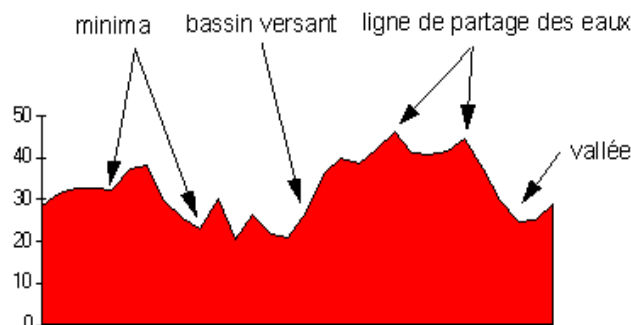
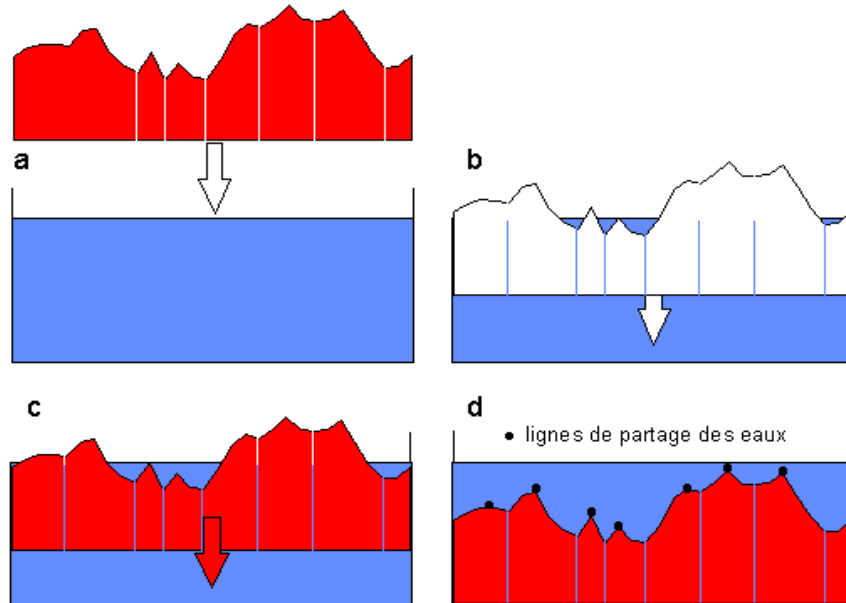


Fig 25. La ligne de partage des eaux comme étant la crête formant la limite entre deux bassins versants

¹ <http://www.tsi.telecom..paristech.fr/pages/enseignement/ressources/beti/lpe2/lpe/lpe.htm>

Chapitre I : Généralités sur le traitement & segmentation d'image

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



Il est à noter que la ligne de partage des eaux n'est pas locale. En particulier, elle n'est pas liée à des structures locales (lignes de crête, escarpement,...). La LPE n'est pas un concept local. Vous ne pouvez pas, avec la seule connaissance locale du voisinage d'un point, répondre à la question : Ce point appartient-il à une ligne de partage des eaux?

Une des difficultés à la mise en œuvre de cette analogie intuitive est qu'elle laisse beaucoup de liberté quant à sa formalisation. Il existe deux classes principales d'implémentation : l'une est basée sur un algorithme récursif d'immersion [L.Vincent.91] et une autre basée sur l'utilisation de fonctions de distances géodésiques [F.Meyer.88].

Notons enfin que l'on n'applique pas la LPE à l'image originale mais à son gradient (morphologique). Cela crée ainsi des lignes de partage des eaux aux points de discontinuités du niveau de gris, ce qui est intéressant pour les applications de segmentation.

Observons maintenant le résultat de la LPE sur une image naturelle simple.

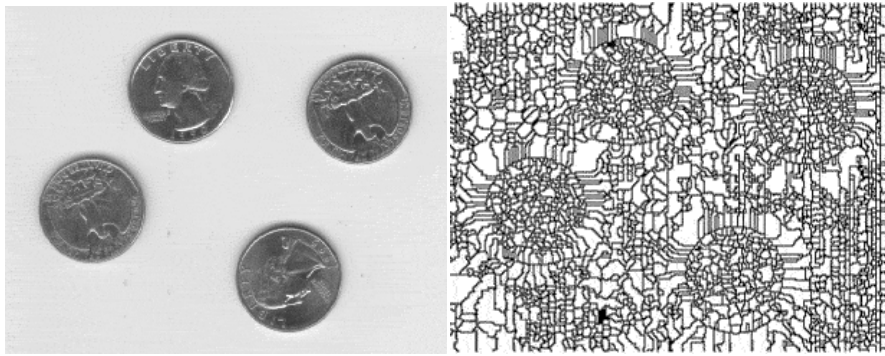


Fig.27 Illustration de la sur segmentation

La LPE du gradient est sur-segmentée, même si comme ici, l'image originale est homogène. Ceci est dû au fait que les images gradient des images naturelles contiennent un grand nombre de minima (dû au bruit du système de prise de vue par exemple ou à variations locales non significatives des niveaux de gris ou encore aux textures des régions). Chaque minimum génère un bassin versant dans la LPE.

Concernant la connexité, on peut envisager l'utilisation de la 4-connexité ou bien de la 8-connexité. Tout dépendra du type d'image et du type de traitement à effectuer. Dans le cadre de la ligne de partage des eaux on utilisera par défaut la 8-connexité.

Il existe plusieurs méthodes qui permettent de pallier à cette sur-segmentation.

- La première consiste à filtrer l'image originale afin de supprimer tous les minima non-significatifs : c'est l'approche filtrage.
- La deuxième consiste à choisir le nombre de minima locaux et donc le nombre de zones que l'on souhaite mettre en évidence grâce à la LPE : c'est l'approche marqueurs (*swamping*).
- Enfin une première LPE peut servir de marqueur pour une seconde, et la zone qu'elle délimite donne lieu à une image mosaïque. Cette image n'est plus construite par pixels, mais comme graphe planaire. Elle est susceptible à son tour d'être traité par LPE, filtrage, *swamping*, versions conditionnelles... C'est un processus itératif qui donne donc lieu à une segmentation hiérarchique.

Chapitre I : Généralités sur le traitement & segmentation d'image

Ceci correspond à un paradigme de segmentation c'est à dire la structure utilisée pour la segmentation à base de LPE. Elle autorise la combinaison des différentes méthodes de prétraitements suscitées.

Un test a été fait sur l'efficacité des prétraitements sur 3 images photographiques : *lena*, et l'image *cell*. Ci dessous les images originales (à gauche) avec le résultat de la LPE (à droite).



Fig.28 image lena

L'image de *lena* est une image riche de détails et de difficultés pour la segmentation : la faible différence de contraste du haut du chapeau, complexité des plumes. Néanmoins certaines de ces caractéristiques sont relativement faciles à extraire comme l'épaule par exemple. C'est une bonne image de test pour la segmentation.

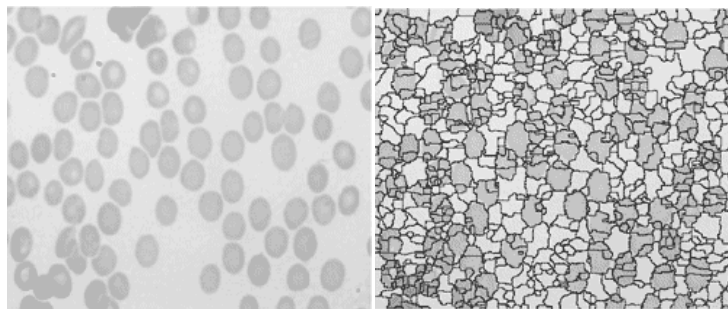


Fig.29 Image cell

Cell est une image typique pour ce genre de segmentation : des cellules avec des niveaux de gris semblables et qui parfois se chevauchent, ce qui rend leur

segmentation difficile. Elle se prête particulièrement aux méthodes utilisant la reconstruction de l'image à partir de marqueurs.

VI-7 Méthodes SLIC: [R.Achanta.11]

Elle adopte le principe de l'approche hybride. La méthode de SLIC superpixel: simple classification itératif linéaire (Simple Linear Iterative Clustering), qui s'adapte à la méthode des k-means pour générer des superpixels d'une manière similaire [C.L.Zitnick&S.Bkang.07] avec deux importantes distinctions:

1) Le nombre des calculs de distance dans l'optimisation est considérablement réduit en limitant l'espace de recherche à une région proportionnelle à la taille de superpixel.

2) Une mesure de distance pondérée combine couleur et spatiale proximité, tout en fournissant simultanément un contrôle sur la taille et la compacité des superpixels.

Bien que remarquablement simple, SLIC est indiqué pour obtenir l'état de l'art respectant les limites d'images sur la référence Berkeley [D.Martin.01], et surpasse les méthodes existantes lorsqu'elle est utilisée pour la segmentation sur le PASCAL [M.Everingham.10]. En outre, elle est plus rapide et plus efficace que les méthodes existantes. En plus de ces avantages quantifiables, SLIC est facile à utiliser, offre une grande flexibilité dans la compacité et nombre de superpixels qu'elle génère, facile à étendre aux dimensions supérieures.



Fig.30: Images segmentées en utilisant SLIC en superpixels de taille 64, 256, et 1024 pixels (environ).

Algorithm 1 SLIC superpixel segmentation

```
/* Initialization */
Initialize cluster centers  $C_k = [l_k, a_k, b_k, x_k, y_k]^T$  by
sampling pixels at regular grid steps  $S$ .
Move cluster centers to the lowest gradient position in a
 $3 \times 3$  neighborhood.
Set label  $l(i) = -1$  for each pixel  $i$ .
Set distance  $d(i) = \infty$  for each pixel  $i$ .
repeat
  /* Assignment */
  for each cluster center  $C_k$  do
    for each pixel  $i$  in a  $2S \times 2S$  region around  $C_k$  do
      Compute the distance  $D$  between  $C_k$  and  $i$ .
      if  $D < d(i)$  then
        set  $d(i) = D$ 
        set  $l(i) = k$ 
      end if
    end for
  end for
  /* Update */
  Compute new cluster centers.
  Compute residual error  $E$ .
until  $E \leq \text{threshold}$ 
```

Fig. 31 *Présentation de l'algorithme SLIC*

CONCLUSION :

Dans ce chapitre nous avons vu deux parties, la première partie nous avons fait un aperçu général sur l'image et sur les différents traitements qui lui sont appliqués afin d'exploiter de façon optimale, les informations qu'elle contient. Dans la seconde partie on s'est focalisé sur la segmentation et ses différentes approches et méthodes.

Les méthodes par histogrammes sont en général rapides à calculer, peu sensibles au bruit. Mais elle n'intègre pas (ou peu) d'information géométrique ou topologique sur les régions. Ce sont des méthodes *globales*, au sens où la décision d'appartenance d'un pixel à une région dépend toujours de l'image entière.

La méthode K-means prend comme représentant de chaque classe son centre de gravité, tout comme la méthode des centres mobiles, mais les centres de gravités sont recalculés à chaque nouvelle affectation d'un pixel à une classe.

Chapitre I : Généralités sur le traitement & segmentation d'image

Les méthodes de croissance de régions qui consistent en une approche ascendante qui sont rapides et faciles à mettre en œuvre mais en revanche leurs algorithmes sont très sensibles au bruit, l'obtention de frontières est non régulière et elles sont peu efficaces dans le cas de dégradés.

Les méthodes de détection de contours, permettent de détecter les transitions entre les régions de l'image. Les détecteurs de contours utilisés peuvent être simples. Toutefois, ils peuvent présenter de fausses détections et ils ne sont généralement pas fermes.

Les méthodes split & merge sont à la fois globales et locales : globales lors de la division, et locales lors de la fusion, moins sensibles au bruit que la croissance et c'est des méthodes parfois complexes d'un point de vue algorithmique : manipulation de lourdes structures de données.

Les méthodes de ligne de partage des eaux nécessitent : la détection des contours, calcul du gradient et la détection des fonds des bassins : minima locaux, elles sont issues de la morphologie mathématique et l'image est considérée comme une surface (terminologie de géographie) : pixels de faible amplitude (faible valeur) = vallée, pixels de forte amplitude = crête. Ligne de partage des eaux = crête formant la limite entre deux bassins versants.

Et à la fin on a vu la méthode SLIC qui permet d'obtenir une image découpée en zones cohérentes, ce qui rend l'image plus facile à analyser. Une telle segmentation est utile notamment pour localiser des objets et des frontières (lignes, courbes, etc.) au sein des images. Plus précisément, les superpixels sont calculés de telle sorte que les pixels appartenant au même superpixel partagent un certain nombre de caractéristiques visuelles. Cette méthode présente de multiples avantages : elle est rapide, robuste, et permet d'obtenir assez aisément des découpages en superpixels de taille quasi-constante sur toute l'image.

Chapitre II :
LA SAILLANCE
VISUELLE

I- INTRODUCTION :

Dans la vie de tous les jours, les humains utilisent la vision beaucoup plus que tous leurs autres sens dans de nombreuses tâches quotidiennes. Une majorité de nos sources d'information est visuelle: lire des livres, regarder la télévision, surfer sur Internet, la localisation spatiale qui oriente et coordonne nos déplacements et fournit une orientation dans nos mouvements.

On estime que près d'un tiers de notre cerveau participe à la transformation des données visuelles. Ainsi, ces tâches nécessitent un système visuel très développé. Lorsque nous créons un système de vision artificielle, il semble intéressant d'apprendre à partir de ce que la nature a longtemps développé et affiné. Cependant, l'architecture de l'ordinateur est assez éloignée de celle du cerveau. Malgré les progrès récents, il n'est pas toujours avec précision de calcul efficace de copier tous les mécanismes de notre système visuel. Une étude, cependant, se révèle être une aide précieuse pour améliorer les algorithmes de traitement de l'information dans la vision par ordinateur. Elle permet de sélectionner les informations les plus pertinentes dans le flux de données auquel nous sommes confrontés: c'est ce qu'on appelle l'attention visuelle. Elle représente l'une des principales caractéristiques du système visuel humain (HVS) pour obtenir d'importantes informations et compacter des scènes naturelles. Depuis le milieu environnant comprenant une quantité excessive de l'information, le mécanisme de l'attention visuelle permet une réduction de la redondance des données qui bénéficie de perception pendant le processus d'attention sélective [L. Itti 00].

II -SYSTEME VISUEL HUMAIN:

L'œil est un dispositif optique biologique classique. Les rayons lumineux passent d'abord à travers la cornée qui agit comme une première lentille, ce qui permet aux rayons lumineux de se concentrer à travers la pupille, le diamètre est fonction de l'ouverture de l'iris. Le diaphragme permet de réguler la quantité de lumière qui passe à travers la lentille pour être projetée sur la rétine. Au centre de la rétine, dans le prolongement de l'axe optique, on trouve la fovéa, où l'on trouve le plus grand nombre de photorécepteurs où notre acuité visuelle est la plus nette. La rétine est la partie

«capteur» de l'œil. Il est composé de deux types de cellules principales. Tout d'abord, nous avons des cellules de photorécepteurs qui comprennent les bâtonnets et les cônes. Leur distribution n'est pas uniforme; les plus éloignés de la fovéa et moins récepteurs sont nombreux. Cônes, principalement utilisés pour la vision de jour, sont concentrés dans la fovéa tandis que les bâtonnets sont principalement utilisés pour la vision de nuit. Deuxièmement, il existe un autre type qui assure la transmission et la transformation de l'information. Le signal des cônes et des bâtonnets est transformé par les cellules horizontales, amacrines et bipolaires avant de traverser les différents types de cellules ganglionnaires. La répartition et les caractéristiques des bâtonnets et des cônes ont un impact significatif sur le fonctionnement de notre système de vision. Dans des conditions diurnes, notre champ de vision est clair et plein de couleurs avec des degrés différents. Considérant que, notre vision périphérique, achromatique, est principalement utilisée dans des conditions de lumière faible. L'image que nous voyons est une résolution variable: précise dans le centre et approximative dans la périphérie [P. Matthieu 10]. Par leur grande sensibilité, les bâtonnets offrent une grande réactivité aux mouvements de la vision périphérique. La capture de mouvement n'est pas très précise (ni dans l'espace, ni dans le temps). Différentes interconnexions entre les cellules forment des champs récepteurs. La plupart sont de type centre-périphérie (centre-Surround), et peuvent être:

- **On centre-off surround** (Fig.1), à gauche, avec une réponse maximale en présence d'une tache blanche entourée de noir.

- **Off center-on surround** (Fig.2) à droite, avec une réponse maximale en présence d'une tache noire entourée de blanc.



Fig.1 : On center-off Surround

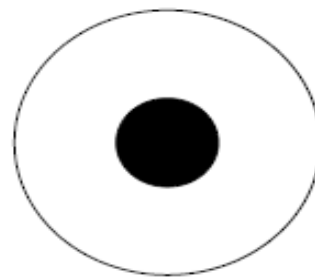


Fig.2 : Off center-on Surround

Les champs récepteurs centre-périphérie [P. Matthieu 10]

Le rôle de ces champs récepteurs est de transformer le signal provenant du photorécepteur, codant l'intensité lumineuse et le contraste en un signal. Cela permet une comparaison du traitement effectué par la rétine avec un algorithme de détection de bord. Ce type de représentation de l'information correspond à un type particulier de codage différentiel: un codage prédictif. Cette méthode est utilisée dans la théorie de l'information pour réduire la quantité de données transférées [K. Sayood. 00]. Sur la base de l'hypothèse que le signal à émettre est essentiellement uniforme et / ou continue, la différence entre les données à coder et la valeur moyenne de son environnement immédiat est généralement plus petit que le nombre de données elle-même.

III- DEFINITION DE LA SAILLANCE :

La saillance contient une idée centrale, mais le détail varie selon le domaine d'application. Un objet est saillant s'il est facilement remarqué. A partir de cette idée intuitive, plusieurs définitions peuvent être proposées, selon le critère qui fait que l'objet va être remarqué. La notion de saillance peut être appliquée au langage comme elle peut être appliquée au visuel (saillance linguistique, saillance visuelle).

La saillance linguistique intervient fortement lors de la lecture d'un texte ou de l'interprétation d'un énoncé en situation de dialogue : mettant en avant un élément du message, elle dirige l'attention du sujet sur cet élément et privilégie sa prise en compte dans le processus d'interprétation, que ce soit au niveau de la détermination du sens ou plus localement lors de la résolution des références.

La saillance visuelle fait largement intervenir la notion de saillance : notre attention s'arrête prioritairement sur les éléments saillants qui ressortent de notre environnement visuel, jusqu'à axer les processus cognitifs sur ces seuls éléments. La saillance naît de causes variées, comme le montre l'exemple suivant : lorsque nous jetons un œil sur une facture suite à des courses dans un supermarché, il arrive qu'un prix élevé attire directement l'attention, sans même nécessiter le parcours détaillé de la liste des articles achetés. La raison peut en être purement visuelle, comme par exemple la présence d'un plus grand nombre de chiffres dans la représentation écrite du prix.

Mais à cette cause visuelle s'ajoute probablement une cause cognitive, liée par exemple à la saillance particulière que prend un prix élevé pour un esprit économe.

D'une manière générale, la saillance visuelle ou linguistique constitue un point de départ pour tout message linguistique. Elle reste néanmoins une notion très floue. La saillance visuelle est liée à la nature et à la disposition d'objets dans une scène, sans être clairement définie par rapport à des propriétés visuelles particulières. La saillance linguistique est liée à la façon dont on présente ou traite un message, et recouvre divers aspects de la structure informationnelle (ou structure communicative), sans être clairement définie par rapport aux composants de cette structure.

Les domaines d'applications de la reconnaissance d'objets sont très vastes : robotique, médecine, sécurité, interaction homme-machine, exploration spatiale, collecte de données. . . . Le principal domaine d'application est l'indexation d'images. On cherche à développer un assistant visuel personnel permettant d'identifier la catégorie à laquelle l'image visualisée appartient et fournissant à l'utilisateur des informations sur cet objet. Une autre application est le développement d'un système qui, à partir d'une série de personnages, actions ou scènes, est capable de retrouver dans une vidéo les endroits où ces éléments apparaissent.

VI- LES FACTEURS DE LA SAILLANCE VISUELLE :

La saillance visuelle, tente la séparation entre facteurs physiques et facteurs cognitifs. Liés aux particularités de la perception visuelle.

VI-1 - Facteurs physiques:

VI-1.1 La saillance liée aux contrastes de luminosité et de couleur:

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

Deuxième facteur que nous proposons, la saillance liée aux variations de perception des couleurs recouvre deux aspects : un premier relatif aux temps de latence, et un second relatif à l'effet spatial des couleurs. Le temps de latence, c'est-à-dire le décalage entre le début de l'excitation et celui de la sensation, varie selon les couleurs. [Y. Baticle 85] donne les chiffres suivants : rouge (22.6 millièmes de seconde) ; vert (37.1 ms) ; gris (43.4 ms) ; bleu (59.8 ms) ; jaune (96.3 ms). Ces chiffres qui constituent une échelle de saillance doivent néanmoins être relativisés : pendant le jour, le maximum de sensibilité est dans le jaune et le rouge, alors que c'est le bleu le plus saillant pendant la nuit. D'autre part, comme le montrent les photographies de [M. Freeman 89]. et les exercices de [J. Itten 85], une couleur appelle sa complémentaire. Un objet qui prend la couleur complémentaire de celle du fond en devient par exemple très saillant. Plus généralement, la saillance de chaque couleur dépend de son association avec d'autres couleurs. [B. Cocula et C. Peyroutet 86] proposent ainsi une hiérarchie de couples de couleurs en fonction de l'importance de leur impact visuel : noir sur blanc ; puis noir sur jaune ; puis rouge sur blanc ; etc. Enfin, [J. Itten 85]. note l'importance de l'effet spatial des couleurs : contrairement à du violet, du jaune sur fond noir a l'air d'avancer et voit ainsi sa saillance augmenter. Ces données nous donnent les bases pour un modèle de la saillance chromatique d'un objet en fonction de la luminosité ambiante, de la couleur dominante du fond et de la couleur dominante de l'objet.

VI-1.2 La saillance liée à une singularité dans un ensemble d'objets:

Il s'agit ici de la saillance due à une mise en relief contextuelle du fait de la présence de plusieurs objets. Compte tenu d'un ensemble d'objets, un élément de cet ensemble est saillant s'il se distingue par une propriété que les autres éléments n'ont pas. Autrement dit, être un singleton est un facteur de saillance. [P. Edmonds 93] donne ainsi l'exemple d'un immeuble a priori saillant par sa grande taille, et qui perd toute saillance lorsqu'il est entouré d'immeubles encore plus grands. En déroulant les types de propriétés d'un objet physique, nous pouvons parler de saillance liée à la catégorie des objets, de saillance liée aux caractéristiques physiques que sont la forme, la couleur, la taille, ou la texture, de saillance liée à l'orientation de l'objet, et de

saillance liée à sa dynamique. Nous considérons qu'il existe une hiérarchie des propriétés qui correspond à l'ordre dans lequel nous les avons citées. La primauté de la catégorie est démontrée par la fréquente nécessité d'utiliser un substantif catégorisant dans une expression référentielle. Si une chaise rouge se distingue dans une scène, on dira plus facilement *la chaise* ou *la chaise rouge* que *l'objet rouge*, le terme *objet* jouant ici le rôle de substantif non catégorisant. En ce qui concerne l'importance relative des propriétés que sont la forme, la couleur et la taille, nous donnons la primauté à la forme car elle correspond souvent à la propriété fondamentale de l'objet. Un exemple de [T. Pattabhiraman 93] le montre clairement dans le cas de comparaisons et de métaphores : dans *les joues de Mary sont comme des pommes, sauf qu'elles sont jaunes*, la comparaison se fait implicitement sur la forme ronde. L'inversion de l'ordre d'importance de la forme et de la couleur aboutit à un exemple déconcertant : *les joues de Mary sont comme des bananes, sauf qu'elles sont rondes*. Entre couleur et taille, des expérimentations avec mesure des temps de réponse montrent que la couleur est traitée plus rapidement que la taille et intervient avant elle dans le processus de perception [Y. Baticle 85]. En génération automatique, [E. Reiter et R. Dale 97] vont dans le même sens en privilégiant la production de l'expression référentielle *le chien noir* plutôt que *le petit chien*, dans une scène confrontant un grand chien blanc et un petit chien noir. En ce qui concerne les saillances liées à l'orientation et à la dynamique, les exemples suivants montrent clairement que le fait d'être un singleton pour l'une de ces propriétés est un facteur de saillance : dans une rangée de chaises, une orientation différente suffit à rendre saillante la chaise incriminée ; dans une scène contenant un objet animé et plusieurs objets inanimés.

VI-1.3 La saillance liée à la cohésion et la structuration de la scène:

Un premier facteur dans cette catégorie est la saillance due à la mise en évidence explicite d'un objet dans la scène. Cette mise en évidence peut être due à un éclairage particulier, par exemple dans le cas d'un comédien éclairé par des spots, ou par un contraste figure sur fond important. Dans le cadre de la communication homme-machine, elle peut également provenir d'une intention ostensive de la part du système, avec l'utilisation d'un rendu visuel particulier (textures transparentes, affichage de la

boîte englobant de l'objet), ou d'une intention ostensive de la part de l'utilisateur, lorsque celui-ci produit un geste de pointage qui rend saillant l'objet désigné.

Un deuxième facteur très proche mais plus implicite est la saillance due à une construction dédiée. C'est le cas du sujet d'une photographie quand celle-ci est composée de manière à valoriser ce sujet, par exemple à l'aide d'une construction en triangle qui, plaçant le sujet au sommet supérieur, le rend saillant [M. Freeman 89].

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

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

Un dernier facteur est la saillance indirecte par transfert visuel de saillance, c'est-à-dire la saillance s'appliquant à un objet qui est relié à un objet fortement saillant, le lien étant dû à une proximité particulière ou découlant des lignes de force qui dirigent le regard de l'observateur compte tenu de la structure de l'image. Ce facteur s'avère difficile à modéliser, d'une part car l'identification de lignes de force est un problème complexe [M. Vettraino-Soulard 93], et d'autre part car le principe de la transitivité tend à rendre saillants tous les objets de la scène et à enlever tout intérêt à la notion de saillance.

VI-2 Facteurs cognitifs:**VI-2.1 La saillance liée à l'intention:**

Un premier facteur cognitif de saillance est ainsi la saillance liée à l'intention, c'est-à-dire à l'influence de la tâche en cours sur la perception. A la suite de son exemple sur les propriétés saillantes d'immeubles, [P. Edmonds 93] souligne que certaines propriétés sont saillantes dans un but précis et pas dans un autre but. Dans son contexte de description d'un itinéraire routier, la propriété de taille est ainsi saillante lorsque la tâche est la désignation d'un immeuble, mais ne l'est pas lorsqu'il s'agit de désigner une intersection de rues. Dans ce sens, un modèle de saillance requiert un modèle de la tâche. La saillance liée à l'intention recouvre également deux autres aspects : les fonctionnalités des objets et leur incongruité compte tenu d'une intention. Les fonctionnalités des objets peuvent être perçues visuellement, et concourir ainsi à la saillance. C'est le cas, pour un utilisateur potentiel, d'un ordinateur allumé dans une pièce contenant d'autres ordinateurs éteints. En ce qui concerne l'incongruité ou l'aspect énigmatique d'un objet compte tenu d'une intention, nous noterons qu'un objet dans une situation incongrue est en infraction avec une règle implicite, culturelle ou fonctionnelle, et acquiert de cette infraction une certaine saillance. C'est le cas d'une chaise renversée ou placée sur une table lorsqu'on a l'intention de s'asseoir, ou d'une chaise placée en plein milieu de la pièce dans laquelle on fait le ménage. De même, une casserole est *a priori* plus saillante dans une salle de bain que dans une cuisine (sauf si elle sert à contenir une fuite d'eau, c'est-à-dire si sa fonction prend le pas sur l'incongruité). Plus particulièrement sur l'aspect énigmatique, nous retiendrons qu'un objet qui n'est pas à sa place compte tenu de la tâche en devient saillant, de même qu'un objet inattendu. Ce dernier exemple se rapproche de l'effet de nouveauté, ou inattendu de [E. Conklin et D. McDonald 82].

VI-2.2 La saillance liée à l'attention:

Un deuxième facteur cognitif est la saillance liée à l'attention visuelle, dans le sens de focalisation à un sous-espace visuel ou de meilleure détection des stimuli du fait

d'une attitude de préparation à la perception. L'attention consiste dans ces deux cas en un traitement prioritaire de données perceptives. L'existence de sous-espaces visuels est montrée par [K. Kessler et al 96]. Dans leur expérimentation, des sujets sont invités à référer linguistiquement à des objets perçus visuellement et séparés en deux groupes perceptifs distincts. Certaines expressions référentielles produites s'avèrent ambiguës si l'on considère la scène complète, mais non ambiguës si l'on se place dans l'un des groupes perceptifs. Ce résultat montre que les sujets ont procédé à une focalisation attentionnelle, qui les a conduits à ignorer le reste de la scène. Un exemple du deuxième cas est celui de personnes qui entrent dans une pièce et cherchent des chaises pour s'asseoir : avec cette attitude de préparation, les chaises deviennent plus saillantes que les autres meubles. Dans ce contexte, l'effet de nouveauté n'est pas un facteur de saillance, ce qui illustre une différence entre intention et l'attention.

VI-2.3 La saillance liée à la mémoire à court terme:

Le troisième facteur cognitif est la saillance liée à la mémoire à court terme, qui concerne les objets récemment traités par le système cognitif visuel, et se rapproche du facteur de vécu récent de [G. Rousselet et M. Fabre-Thorpe 03]. Le principal critère pour une modélisation de ces aspects est la capacité limitée à sept éléments, capacité revue à cinq ou quatre éléments dans des travaux plus récents (cités par [G. Rousselet et M. Fabre-Thorpe 03]). Reste à déterminer quels sont les quatre ou cinq éléments visuels qui seront retenus. La récence de traitement n'est en effet par le seul critère, comme le montre [Z. Pylyshyn. 01] avec la notion d'index visuel. Un index est une sorte d'étiquette ou de fichier. Lié à un élément visuel et destiné à le repérer rapidement (sans le décrire), il ne s'applique qu'aux éléments saillants d'une scène, ceux qui seront traités et retenus de manière prioritaire [Z. Pylyshyn. 01] décrit le fonctionnement de ces index, et nous pouvons considérer à sa suite que les quatre ou cinq derniers objets repérés par un index sont cognitivement plus saillants que les autres objets visuels.

VI-2.4 La saillance liée à la mémoire à long terme:

Quatrième facteur, la saillance liée à la mémoire à long terme peut se décomposer en deux aspects, premièrement la familiarité visuelle individuelle, et deuxièmement la familiarité visuelle culturelle. Un exemple de familiarité individuelle est celui de l'infographiste habitué à travailler sur des images et à corriger leurs couleurs : pour lui, un contraste entre deux teintes de rouge sera saillant, alors que le même contraste pourra ne pas être perçu par un autre individu. Plus généralement, nous acquérons tous nos propres sensibilités à certaines couleurs, nos propres critères de saillance. Dans le cadre de la communication homme-machine, un modèle de saillance requiert dans ce sens un modèle de l'utilisateur. Un exemple de familiarité culturelle est le fait que, dans notre vie en société, nous avons l'habitude de communiquer avec nos semblables. Il est ainsi possible de classer selon leur saillance les participants d'une situation visuelle : la présence d'un être humain dans notre champ de vision est saillante avant tout (surtout s'il s'agit de quelqu'un que l'on connaît) ; vient ensuite la saillance d'un animal ; puis celle des objets inanimés. Un autre exemple relatif à une communauté est celui des couleurs du drapeau national. [Rastier F] note que les couleurs de son propre pays sont culturellement saillantes pour un individu. [M. Segall et al 66] montre de manière détaillée comment la perception visuelle varie selon les cultures. Avec comme base d'expérimentation certaines illusions visuelles classiques, ils distinguent par exemple les cultures « circulaires » habituées aux courbes et peu affectées par les illusions mettant en jeu des droites, et les cultures comme la nôtre où les lignes droites et parallèles sont fréquentes (rues, immeubles, etc.). Plus généralement, l'appartenance d'un individu à une communauté, quelle qu'elle soit, entraîne la possibilité de facteurs spécifiques de saillance.

VI-2.5 La saillance liée à l'affect et aux émotions:

Tout stimulus visuel peut en effet provoquer chez un individu particulier certaines émotions. [T. Pattabhiraman 93] parle ainsi de vivacité pour décrire la capacité d'un stimulus à évoquer des représentations émotives, par exemple, la perception du visage d'un interlocuteur peut conduire à l'identification immédiate de ce que ressent cet

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

V- METHODES DE SAILLANCE: BOTTOM-UP ET TOP-DOWN

Le système de l'attention humaine est influencé par deux sources d'information: bottom-up et top-down. Un modèle artificiel de l'attention visuelle devrait donc tenir compte de ces deux éléments. Cependant, le modèle top-down est difficile à comprendre car il dépend, par définition, sur le contexte. Ceci est certainement la raison pour laquelle la grande majorité des modèles à la fin des années 1990 et au début des années 2000 ne portait que sur la composante ascendante [J. Tsotsos 05]. Notez, cependant, quelques exceptions avec des modèles bottom-up et top-down très proches de la théorie [R. Desimone 95], mais aussi des approches plus de calcul [S. Ahmad 92] et [R. Milanese 94], où la mise en œuvre du mécanisme Top-down consiste soit à la capacité globale [S. Ahmad 92] ou locale [C. Mozer. 98] de changement de poids des cartes, soit une carte calcul descendante d'un système de reconnaissance d'objets [R. Milanese 94]. L'ajout d'un composant de haut en bas permet de créer un système d'attention supervisé, capables de s'adapter à un contexte spécifique d'utilisation. Ainsi, la plupart des modèles récents proposent d'intégrer les deux sources d'attention. Le modèle Top-Down peut être pris en compte lors de l'utilisation d'un des changements globaux ou locaux en poids de chaque caractéristique (intensité, couleur, orientation, etc.) en fonction des propriétés de l'objet à reconnaître. Cela peut être fait soit en fournissant directement la carte correspondante à la cible, [F. Hamker. 05], ou par un processus d'apprentissage pour calculer le poids séparant la cible de l'arrière-plan, [S. Choi 06]. Dans [S. Dong 06] les auteurs utilisent la couleur de la peau pour générer une carte de haut en bas qui résulte un système de reconnaissance de visage alors que dans [K. Rapantzikos 03], les auteurs utilisent une carte de haut en bas qui résulte un système de détection de la couleur de la peau.

IV - MODELES DE CARTE DE SAILLANCE:

Ces modèles sont dans la continuité de la théorie de l'intégration de [A. Treisman 80]. Elle suppose qu'il existe une carte maîtresse (la carte de saillance), une structure bidimensionnelle stagiaire qui représente le champ visuel externe. Ce mode de représentation est très populaire en raison de la simplicité de la représentation et de l'efficacité de calcul, mais des études récentes n'ont pas réussi à prouver l'unicité de la représentation de saillance dans notre cerveau.

[P. Matthieu 10] propose une taxonomie de ces méthodes dans cinq familles que nous citons ci-dessous. Nous extrapolons cette taxonomie en ajoutant deux familles: fréquentielle et graphique.

IV-1 hiérarchique:

Ce type d'algorithme construit à partir d'une image initiale, une hiérarchie des différentes cartes de fond, qui va progressivement être combinés pour obtenir une représentation unique centrale: la carte de saillance. L'un des modèles les plus influents de cette classe d'algorithmes est le modèle de Laurent [L. Itti 00] qui est l'un des premiers modèles de calcul de l'attention qui a été améliorée constamment depuis sa création.

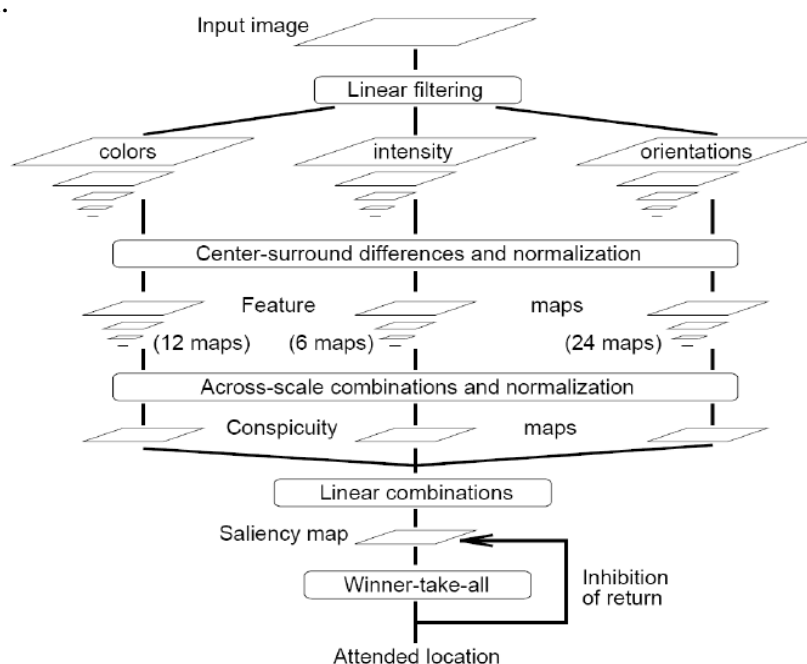


Fig 2 : Structure du modèle d'attention visuelle du [L. Itti 00]

L'architecture de ce modèle de bas en haut (*fig.3*) est basée sur les étapes suivantes.

L'image source est décomposée en différents canaux de perception (dans la version de base de l'algorithme: l'intensité, la couleur et orientation).

Puis, une représentation multi-échelle est construite à partir de ces différents canaux. Une opération de filtrage centre et la périphérie est alors effectuée en soustrayant certains niveaux de pyramides multi-échelles afin d'obtenir différentes cartes de fonction. Ceux-ci sont ensuite normalisés par N opérateurs de renforcement des cartes contenant seulement un petit nombre de pic d'activité (des zones en relief) et ensuite additionnées pour obtenir trois cartes différentes (à grande visibilité) d'intensité, couleur, orientation. Ces cartes sont également normalisées avec l'opérateur N, puis additionnées pour obtenir une carte de saillance. La génération de différents regards est assurée par un réseau de Winner Takes All (WTA), sélectionnant la zone d'activité maximale dans la carte de saillance, couplée avec une inhibition de mécanisme de retour, désactivant temporairement les zones déjà visitées afin que le centre d'attention ne revienne immédiatement. Il existe deux modèles particuliers provenant de [L. Itti 00] parce qu'ils apportent des améliorations intéressantes et complémentaires: d'abord, le modèle d'Olivier [O. Le Meur 05] pousse l'inspiration biologique bien au-delà du modèle original. Sans l'aide d'approche neuronale comme ce qui est utilisé dans les modèles distribués, il modélise finement beaucoup de phénomènes psycho-visuel: les fonctions de sensibilité au contraste, complexe décomposition en canaux perceptifs, etc. Cette approche montre qu'il est possible d'obtenir une très belle modélisation avec une approche plutôt macroscopique (le comportement des neurones individuels dans le cortex visuel n'est pas abordé du tout). Certains modèles hiérarchiques sont basés sur un objet plutôt que l'attention spatiale. Parmi ces modèles peuvent être mentionnés proposées par [F. Liu 06] et [T. Geerinck 09]. Ces études diffèrent sur la façon dont des objets ou des régions segments (segmentation couleur, le bassin versant de l'intensité du gradient, etc.), mais tous ont le même but: fournir un accent sur les régions, plutôt que des points.

IV-2 Théorie de l'information:

La théorie des modèles d'information supposent que le cerveau humain utilise des mécanismes volontaires pour maximiser la quantité d'information acquise. Estimée localement, elle peut alors être utilisée pour définir une saillance dans une image. Différentes approches de calcul de l'estimation de la quantité d'informations est possible. [S. Gilles 96] fournit l'explication d'une saillance en termes de complexité locale qui peut être mesurée par l'entropie de Shannon sur des attributs locaux d'une image. [T. Kadir. 01] prend cette définition et étend le modèle en utilisant le maximum d'entropie pour déterminer l'échelle des éléments saillants dans une analyse multi-échelle. [B. Bruce. 03] propose d'utiliser une mesure de bonne information (auto-information) pour construire les opérateurs de filtrage non-linéaires, utilisés pour normaliser les cartes de singularité avant leur fusion, à l'architecture proche de celle proposée par [L. Itti 00]. Il réutilise ce principe dans son travail avec [J. Tsotsos 05], combinant l'analyse indépendante de composant (ICA) [Parc 02] et la mesure d'une information appropriée dans le but d'obtenir une estimation de la saillance dans une image. [M. Mancas 07] propose une approche globale de la saillance basée sur l'information appropriée.

IV-3 Connexionniste:

Les modèles de cette catégorie utilisent principalement les réseaux de neurones comme mécanisme attentionnel. Bien que basé sur la concurrence, ces modèles font partie des algorithmes centralisés car ils utilisent une carte de saillance. L'un des premiers modèles connexionnistes en utilisant une représentation centralisée de saillance est le modèle « VISITE 6 » de [S. Ahmad 92]. Ce traitement se divise en trois attentionnel réseaux indépendants, connectés à une mémoire de travail (*fig. 4*). Le réseau de transmission sélective supprime toute activité ne se trouve pas dans une région spécifique (la concentration de l'attention). Le réseau prioritaire sélectionne les zones d'intérêt en utilisant bottom-up et / ou des informations top-down. Ces informations doivent être fournies via une carte centrale appelée priorité (carte de priorité). Le réseau de contrôle fait le lien entre les deux réseaux ci-dessus et détermine l'évolution de la mise au point de l'attention.

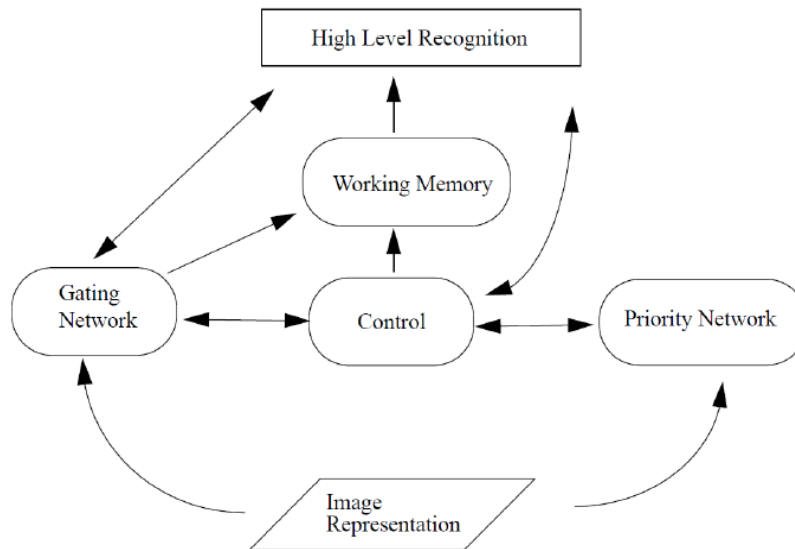


Fig.3 : Le Modèle connexionniste [S. Ahmad 92]

IV-4 Algorithmique:

[M. Lopez 06] propose un modèle en utilisant uniquement la forme et les caractéristiques de mouvement afin d'améliorer la segmentation de la vidéo. Le modèle attentionnel proposé est construit selon l'application cible et est donc difficile de généraliser. Il montre, cependant, que d'un modèle attentionnel peut améliorer considérablement les tâches les tâches de certaines tâches de vision de l'ordinateur. Le modèle [F. Orabona .08] est basé sur le concept proto-objet. L'image est d'abord segmentée en taches de couleur uniforme. Une carte de saillance est ensuite calculée par l'exploitation d'un centre-périphérie filtrage entre chaque forme et ses environs. Une approche similaire est utilisée par [M. Aziz 09]. L'image segmentée en régions de couleurs uniformes est utilisée pour calculer des caractéristiques différentes (contraste de couleur, la taille, la symétrie, l'orientation et l'excentricité). Une carte de saillance globale est ensuite calculée en fonction de la rareté locale de chaque caractère dans chaque région de l'image. Notez que ce modèle est l'un des rares modèles de l'attention avec [G. Sela et D. Marlin 97] pour tenir compte de la symétrie.

Récemment, [M.Cheng et al.13] proposent un algorithme basé sur l'extraction régionale de contraste de saillance avec au milieu des indices des niveaux qui évaluent simultanément les différences de contraste global et la cohérence spatiale. Dans [F. Liu 06], trois cartes de fonction basée sur les caractéristiques locales, régionales, ainsi que

mondiales sont développées et intégrées à un modèle formé de champ aléatoire conditionnel. [J. Feng et al 11] présentent une méthode basée « segment » dans lequel la saillance visuelle est mesurée par le coût de la composition d'un patch de fenêtre avec les parties restantes de l'image. [L. Marchesotti et al. 10] exploitent la similitude visuelle entre images pour calculer la carte de saillance. Étant donné une image de test, ils récupèrent d'abord un certain nombre de base de données les plus similaires et puis apprennent une simple classification avec les noyaux pour séparer les régions saillantes (au premier plan) et non saillantes (arrière-plan). Cependant, cette méthode entraîne inévitablement des charges de calcul importantes en raison de la recherche d'image visuelle. Récemment, [M. Ran 13] propose un algorithme qui intègre motif et la couleur distinction d'une manière unique. Le motif de distinction est déterminé en considérant les statistiques internes des correctifs dans l'image.

IV-5 Graphique:

[Y. Chuan 13] classent la similitude des éléments d'image (pixels ou régions) avec des propriétés (indices) d'avant plan ou d'arrière-plan via un graphe basé sur le manifold ranking. La saillance des éléments d'image est définie en fonction de leur pertinence pour le graphe de requêtes données avec des superpixels comme des nœuds, qui sont classés en fonction de la similitude de fond et les requêtes de premier plan. [L. Wang 11] analyse plusieurs propriétés dans un cadre de minimisation d'énergie unifiée et utilise un modèle de saillance à base de graphes [J. Harel 06] pour détecter des objets saillants. Dans [Y. Lu et al 11] développent un modèle de graphe hiérarchique et utilisent le contexte de la concavité pour calculer les poids entre les nœuds, à partir de laquelle le graphique est bi-partitionné pour la détection d'objet saillant. [V. Gopalakrishnan et al 10] formulent le problème de détection d'objet comme une segmentation binaire ou une tâche d'étiquetage sur un graphique. La plupart des requêtes d'arrière plan saillantes sont identifiés par le comportement des marches aléatoires sur un graphe complet et un graphe de k-régulier. Puis, une technique semi-supervisée d'apprentissage est utilisée pour déduire les étiquettes binaires des nœuds non étiquetés.

CONCLUSION :

Dans ce chapitre on a vu la notion de saillance visuelle qui varie selon le domaine d'application. Et donc on dira qu'un objet est saillant s'il est facilement remarqué, à partir de cette idée, plusieurs peuvent être proposées, selon le critère qui fait que l'objet va être remarqué.

Dans la saillance visuelle il existe deux notions ; la première notion est appelée saillance de recherche, c'est le résultat d'une recherche d'un objet volontaire, et est conditionnée par le processus différents, notamment l'attention focalisée et la vigilance. La seconde notion est appelée saillance attentionnelle, qui est involontaire et principalement liée aux propriétés du stimulus.

Dans notre étude nous allons nous concentrer sur la saillance attentionnelle qui se repose sur les mécanismes élémentaires de la vision, elle s'appuie sur un modèle de traitement de l'information qui mime le fonctionnement du système visuel humain ainsi qu'une modélisation par analyse d'image, cette dernière est le passage du niveau des pixels au niveau des objets ce qui implique, concrètement, une phase de segmentation des objets dans l'image, qu'on a étudié dans le premier chapitre.

Chapitre III
CHOIX DE LA METHODE
&
MSE EN ŒUVRE

***«DETECTION DE LA SAILLANCE VIA
UNE METHODE BASEE SUR UN
MULTI-GRAPHS »***

I- PRESENTATION DE L'ENVIRONNEMENT DE TRAVAIL :**I-1 Présentation de MATLAB :**

MATLAB (MATrix LABoratory) est un logiciel interactif basé sur le calcul matriciel. Il est utilisé dans les calculs scientifiques et les problèmes d'ingénierie parce qu'il permet de résoudre des problèmes numériques complexes en moins de temps requis par les langages de programmation courant, et ce grâce à une multitude de fonctions intégrées et à plusieurs programmes outils testés et regroupés selon usage dans des dossiers appelés boîtes à outils ou "toolbox".

La version de MATLAB utilisée dans notre application est la suivante :

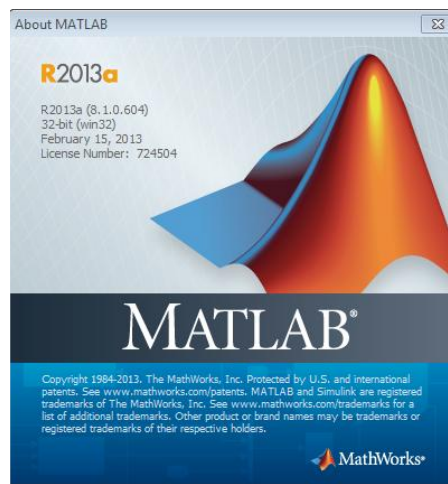


Fig.1 lustration de la version de MATALB (version 8.1)

Au démarrage de MATLAB sur un PC, l'interface se présente comme suit :

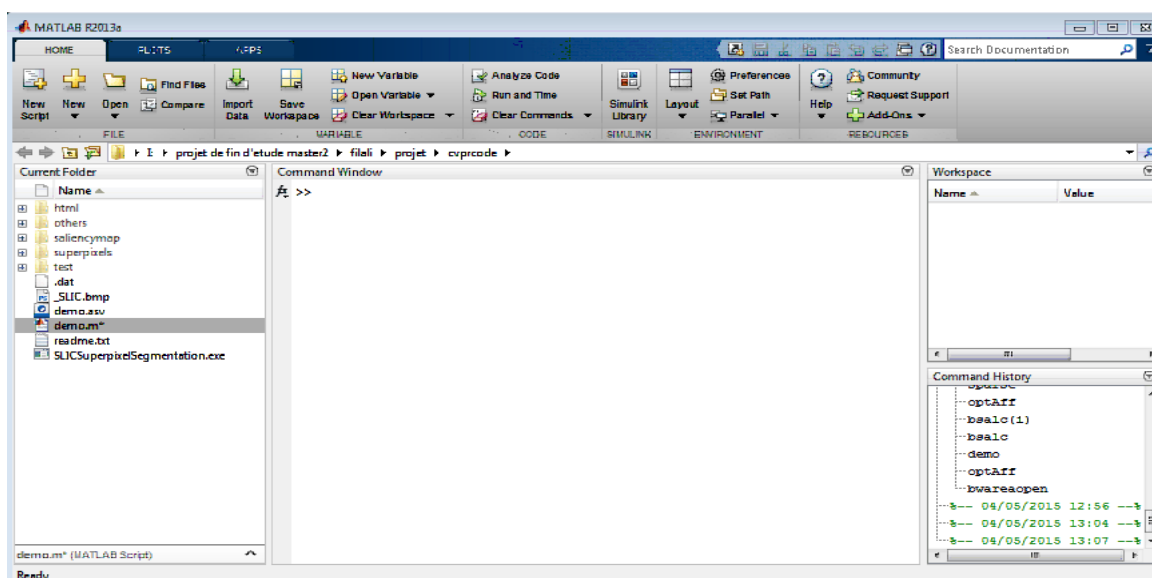


Fig.2 : fenêtre principale de MATALB

La fenêtre principale de MATLAB contient trois fenêtres secondaires pouvant être déplacées ou fermées. Au centre la fenêtre **Command Window** permet à la fois d'afficher les résultats et d'introduire les commandes ligne par ligne. Le symbole ($>>$) indique l'attente d'une commande.

A droite, sont imbriquées en volets les fenêtres Workspace, Command History. et à gauche la fenêtre Current Folder.

- **Workspace** permet d'afficher les variables utilisées avec leurs tailles.
- **Current Folder** affiche le chemin d'accès ou le répertoire en cours avec tous les fichiers et les sous répertoires.
- **Command History** affiche les commandes ayant été saisies.

Les déclarations et les commandes peuvent aussi être introduites sous forme d'un *script* dans un fichier texte d'extension ".m". MATLAB est équipé d'un éditeur de texte permettant de saisir les fichiers script. La commande **edit** ouvre l'éditeur et charge le fichier prog1.m s'il existe, sinon l'éditeur s'ouvre sur un fichier vide. La figure suivante montre l'éditeur de MATLAB ouvert sur le fichier **demo.m** en cours de création.

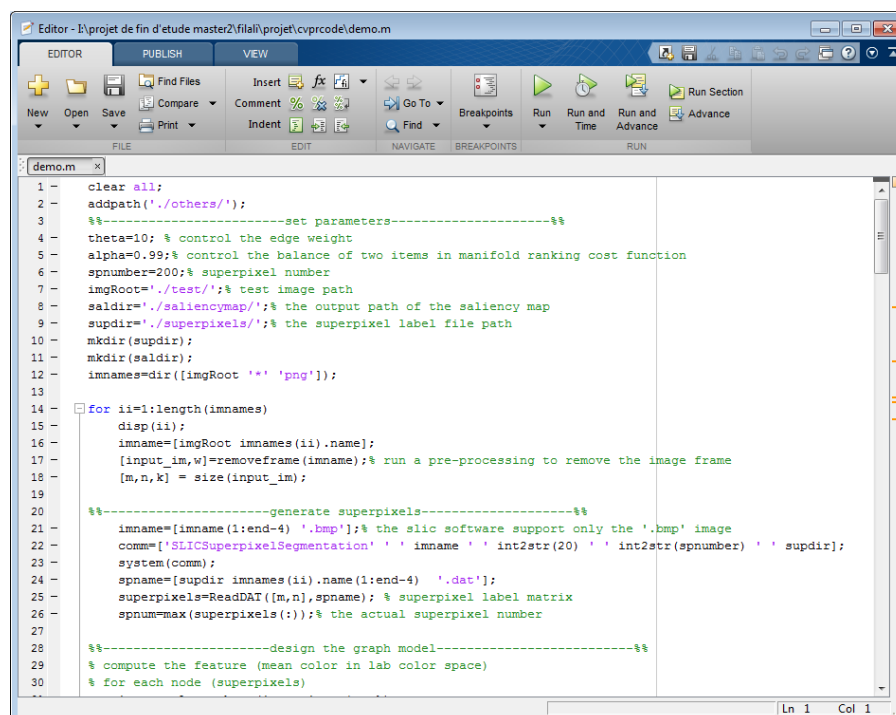


Fig.3 : Editeur de texte MATAB

L'exécution du script (les commandes une après une) se fait à l'aide du bouton **Save** and **Run** sélectionné dans (*la fig.3*) ou avec le menu **debug/Save** and **Run** ou bien, simplement, en appuyant sur la touche fonction **F5**. Les résultats sont affichés dans la fenêtre des commandes.

Une ligne de commande peut contenir plusieurs instructions séparées par des virgules (,) ou par des points-virgules (;). Le résultat d'une instruction suivie par un point-virgule ne sera pas affiché. Le caractère pourcent (%) est utilisé pour les commentaires.

MATLAB fonctionne essentiellement avec des matrices multidimensionnelles. Les composantes peuvent être réelles, complexes ou symboliques. Ainsi, les scalaires sont représentés par des matrices 1×1 et les vecteurs à n composantes par des matrices $1 \times n$ pour les vecteurs lignes et $n \times 1$ pour les vecteurs colonnes.

Il est possible d'introduire une matrice dans MATLAB en saisissant explicitement la liste de ses éléments dans la fenêtre des commandes ou en la chargeant à partir d'un fichier. Les colonnes sont séparées par des virgules (,) ou des espaces et les lignes par des points-virgules (;) ou des sauts de ligne (RETURN).

II- PRESENTATION DE LA METHODE :

II-1. Introduction :

La tâche de détection de la saillance est d'identifier la plus importante partie informative d'une scène. Elle a été appliquée à de nombreux problèmes de vision, y compris la segmentation d'images [S. Goferman et al.10], reconnaissance d'objets [U. Rutishauser et al.], la compression d'image [L. Itti.04], le contenu en fonction de récupération d'image [T. Chen et al.], pour n'en nommer que quelques-uns.

Les méthodes de saillance en général peuvent être classées comme des approches bottom-up (ascendantes) ou top-down (descendantes). Les méthodes ascendantes [R. Achanta et al.09] sont pilotées par les données et pré-attentifs, alors que les méthodes descendantes [T. Liu et al.11] sont des tâches qui impliquent l'apprentissage supervisé avec étiquetage de classe. Nous notons que les modèles de la

saillances ont été développés pour la prédiction de fixation d'œil et la détection d'objet saillant. Les anciens se concentraient sur l'identification de quelques emplacements de fixation humaine sur des images naturelles, ce qui est important pour la compréhension de l'attention humaine. Ce dernier doit précisément détecter où l'objet saillant devrait être, ce qui est utile pour beaucoup de tâches de vision de haut niveau. Dans ce chapitre, nous nous concentrons sur les tâches de détection d'objets saillants dans des images en couleur.

Nous générons des superpixels pour chaque image avec la méthode « *SLICS* » [R.Achanta.11] avec deux granularités.

Pour chaque image, on construit un graphe en boucle fermée où chaque nœud est un superpixel. Nous modélisons la détection de la saillance comme un problème multi-graphes basé sur la méthode Manifold Ranking et proposant un schéma en deux étapes pour l'étiquetage de graphe : dans la première étape, on exploite la limite d'avant [L. Grady et al.11],[V. Lempitsky et al.09] en utilisant les nœuds de chaque côté d'arrière-plan de l'image en tant que requêtes étiquetées. A partir de chaque résultat marqué, nous calculons la saillance de nœuds en fonction de leurs pertinences, par exemple : la couleur qui fait transformation dans l'espace de couleur LAB, l'entropie et la variance.

Les quatre cartes marquées sont ensuite intégrées pour générer une carte de saillance de première étape. Dans la deuxième étape, nous appliquons la segmentation binaire sur la carte de saillance résultat de la première étape, et nous prenons les nœuds de l'avant plan étiquetés comme des requêtes saillantes se qui est montré par la **Fig.4**

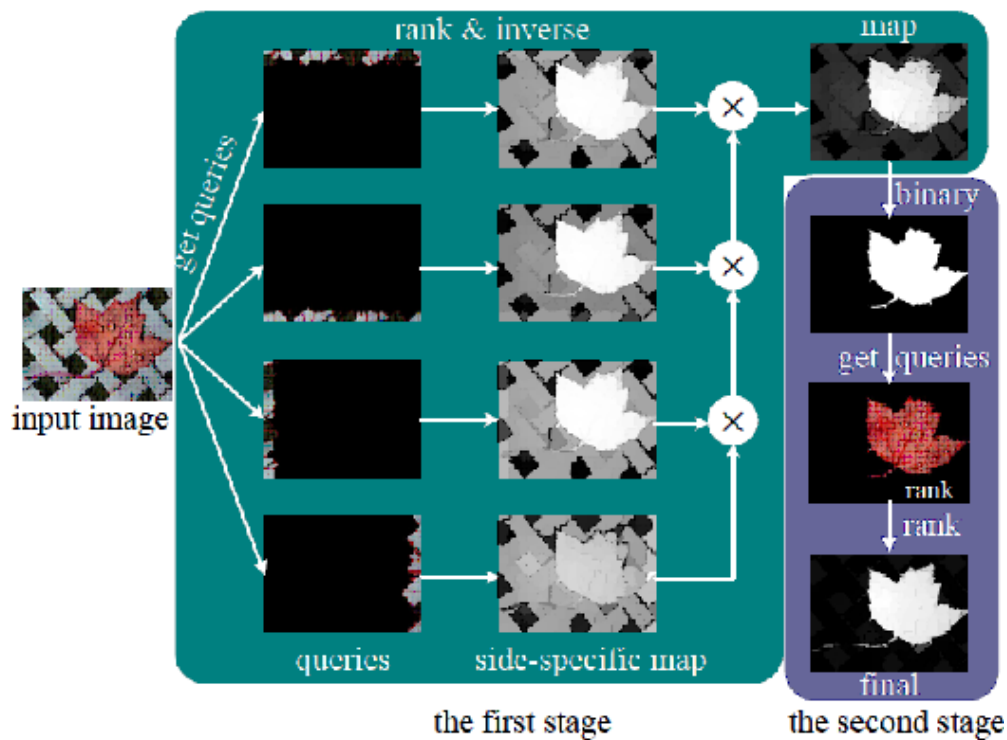


Fig.4 Génération des quatre cartes de saillance

La saillance de chaque nœud est calculée sur la base de sa pertinence aux requêtes de l'avant-plan. Pour capturer pleinement la structure d'informations du graphe intrinsèque et incorporer les propriétés de regroupement locales dans l'étiquetage de graphe, nous utilisons des techniques de ranking différentes pour apprentissage d'une fonction de Ranking, qui est essentielle pour l'établissement d'une matrice d'affinité optimale.

Dans ce travail, toutes les requêtes d'arrières plan et de l'avant plan peuvent être facilement générée via les requêtes antérieurs et le classement d'arrière plan. Comme notre modèle incorpore des propriétés de groupement locales extraites de l'image entière, l'algorithme proposé génère des objets saillants et met uniformément en évidence les régions saillantes entières. Des résultats expérimentaux utilisant une large référence de données qui montrent que l'algorithme de détection de saillance proposé retourne des résultats plus favorables en comparaison avec d'autres méthodes de l'état de l'art.

II-2 Multi- graphes basé sur Manifold Ranking :

Le problème de classement basé sur les graphes est décrit comme suit: étant donné un nœud comme une requête, les nœuds restants sont classés en fonction de leurs pertinences à la requête donnée.

Méthode de Ranking :

Dans [D.Zhou et al.03], une méthode de Ranking qui exploite les diverses structures intrinsèques des données pour le graphe étiqueté est proposé. Étant donné un ensemble de données $X = \{x_1, \dots, x_l, x_{l+1}, \dots, x_n\} \in \mathbb{R}^{m \times n}$, certains points de données sont des requêtes étiquetées et le reste doivent être classés en fonction de leurs pertinences aux requêtes. Soit $f: X \rightarrow \mathbb{R}^n$ représentent une fonction de ranking qui assigne une valeur de ranking f_i à chaque point x_i , et f peut être considéré comme un vecteur $f = [f_1, \dots, f_n]^T$. Soit $y = [y_1, y_2, \dots, y_n]^T$ désigner un vecteur d'indication, dans lequel $y_i = 1$ si x_i est une requête, et $y_i = 0$ sinon. Ensuite, nous définissons un graphe $G = (V, E)$ sur l'ensemble de données, où les nœuds V sont l'ensemble de données X et les bords E sont pondérés par une affinité matrice $\mathbf{W} = [w_{ij}]_{n \times n}$. étant donné un graphe G , la matrice de degré est $\mathbf{D} = \text{diag} \{d_{11}, \dots, d_{nn}\}$, où $d_{ii} = \sum_j w_{ij}$. Similaire au PageRank et aux algorithmes spectraux de clustering, le ranking optimal de requêtes est calculé en résolvant le problème d'optimisation suivant:

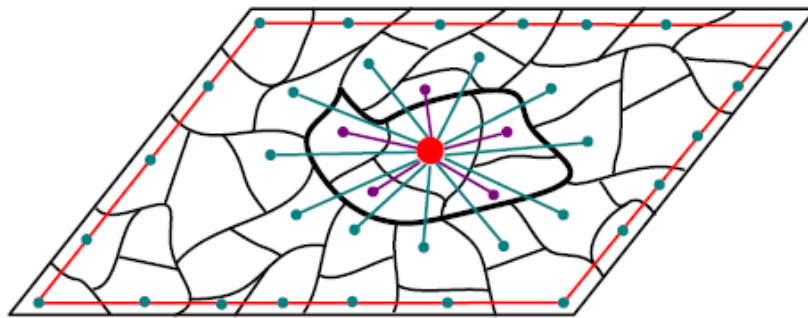


Fig 5. Notre modèle de graphe. La ligne rouge le long des quatre côtés indique que tous les nœuds de frontière sont reliés les uns aux autres.

$$\mathbf{f}^* = \arg \min \frac{1}{2} \left(\sum_{i,j=1}^n w_{ij} \left\| \frac{f_i}{\sqrt{d_{ii}}} - \frac{f_j}{\sqrt{d_{jj}}} \right\|^2 + \mu \sum_{i=1}^n \|f_i - y_i\|^2 \right), \quad (1)$$

Où le paramètre μ contrôle l'équilibre de la contrainte de lissage (le premier terme) et la contrainte de montage (le second terme). C'est une bonne fonction de classement qui ne devrait pas changer trop entre des points voisins (des contraintes de lissage) et ne devrait pas être trop différente de la requête initiale (contrainte de montage). La solution minimale est calculée en définissant la dérivée de la fonction ci-dessus pour être égal à zéro. La fonction de classement résulté peut être écrite comme:

$$f^* = (I - \alpha S)^{-1}y, \quad (2)$$

Où I est une matrice d'identité, $\alpha = 1 / (1 + \mu)$ et S est une matrice Laplacienne normalisée, $S = D^{-1/2} W D^{-1/2}$.

L'algorithme de Ranking [D. Zhou et al.04] est tiré des travaux de l'apprentissage semi-supervisé pour la méthode de Manifold Ranking, où seulement des exemples positifs ou négatifs sont requis. Nous pouvons obtenir une autre fonction de classement en utilisant la matrice Laplacien non normalisée dans *l'équ.2*:

$$f^* = (D - \alpha W)^{-1}y. \quad (3)$$

Ensuite on mesure la saillance de chaque nœud (pixel) : la saillance de chaque nœud est définie comme son score de ranking calculé par *Eq.3* qui est réécrite comme $f^* = Ay$ pour faciliter l'analyse. La matrice A peut être considérée comme une matrice d'affinité optimale apprise qui est égale à $(D - \alpha W)^{-1}$. Le score de classement $f^*(i)$ du nœud i -ème est le produit intérieur de la ligne i -ème de A et y . Parce que y est un vecteur d'indicateur binaire, $f^*(i)$ peut également être considéré comme la somme des pertinences du nœud i -ème à toutes les requêtes. Dans les problèmes de classement conventionnel, les requêtes sont étiquetées manuellement avec le masque binaire. Cependant, comme les requêtes pour la détection de saillance sont sélectionnés par l'algorithme proposé, certains d'entre eux peuvent être incorrect. Ainsi, nous avons besoin de calculer un degré d'assurance (c'est à dire, la valeur de la saillance) pour chaque requête, qui est défini comme son score de classement classé selon les autres requêtes (sauf lui-même). À cette fin, nous mettons les éléments

diagonaux de la matrice A à 0 lors du calcul du classement marqué par *l'équ.3*. Nous notons que ce processus apparemment insignifiant à de grands effets sur les résultats finaux. Si nous calculons la saillance de chaque requête sans définir les éléments diagonaux de la matrice A à 0, sa valeur de classement dans f^* contiendra la pertinence de cette requête à lui-même, qui n'a pas de sens et souvent anormalement grande, afin d'affaiblir gravement les contributions des autres requêtes au score de classement. Enfin, nous mesurons la saillance des noeuds en utilisant le score de classement normalisée \bar{f}^* lorsque les requêtes saillantes sont donnés, et en utilisant $1-\bar{f}^*$ quand les requêtes d'arrières plan sont données.

II-3. La Construction de graphe :

Nous construisons un graphe à une seule couche $G = (V, E)$, où V est un ensemble de noeuds et E est un ensemble de bords non orientés. Dans ce travail, chaque noeud est un superpixel généré par l'algorithme SLIC [R. Achanta et al.11]. Comme les noeuds voisins vont probablement partager l'apparence similaire et la valeur de la saillance, nous utilisons un graphe de k -régulier pour exploiter la relation spatiale.

Tout d'abord, chaque noeud est non seulement connecté à ces noeuds voisins, mais également connecté aux noeuds partageant les frontières communes à son noeud voisin. En prolongeant la portée de la connexion de noeud avec le même degré de k , nous utilisons efficacement des propriétés de lissage locales. Deuxièmement, nous imposons que les noeuds des quatre côtés de l'image soient connecté, c'est à dire, n'importe quelles paires de noeuds de limites sont considérées adjacentes. Ainsi, on note le graphe comme un graphe en boucle fermée. Cette contrainte de boucle fermée améliore significativement la performance de la méthode proposée comme il a tendance à réduire la distance géodésique de superpixels similaires, ce qui améliore les résultats du classement.

Nous notons que ces contraintes fonctionnent bien lorsque les objets saillants apparaissent près des limites d'image ou certaines des régions d'arrière-plan ne sont pas identiques.

Avec les contraintes sur les bords, il est clair que le graphe construit est peu connecté. C'est-à-dire, la plupart des éléments de la matrice d'affinité \mathbf{W} sont nuls (à zéro). Dans ce travail, le poids entre deux nœuds est définie par :

$$w_{ij} = e^{-\left(\frac{\|c_i - c_j\|}{\sigma^2} + \frac{\|e_i - e_j\|}{\omega^2} + \frac{\|v_i - v_j\|}{\partial^2}\right)} \quad i, j \in V, \quad (4)$$

Où c_i représente la moyenne de colorimétries des pixels appartenant au superpixel i correspondant dans l'espace de couleur CIE LAB, e_i représente la moyenne de l'entropie des pixels appartenant au superpixel i , v_i représente la moyenne de la variance des pixels appartenant au superpixel i et σ , ∂ , ω sont des constantes qui contrôlent la force du poids. Les poids sont calculés en fonction de la distance dans l'espace de couleur comme il a été montré pour être efficace dans la détection de saillance.

En construisant le graphe avec les nœuds, la matrice inverse $(\mathbf{D} - \alpha\mathbf{W})^{-1}$ dans l'équ.3 peut être considérée comme une matrice d'affinité complète, c'est à dire, il existe une valeur de pertinence non nulle entre n'importe quelle paire de nœuds sur le graphe. Cette matrice capture naturellement des informations de relation spatiales. Autrement dit, la pertinence entre les nœuds est augmentée lorsque leur distance spatiale est diminuée, ce qui est une propriété importante pour la détection de saillance.

II-4 Détection de saillance en deux étapes :

Dans cette section, nous détaillons le plan proposé en deux étapes pour la détection de saillance en utilisant le ranking avec les requêtes d'arrière plan et d'avant plan.

II-4-1 Classement avec des requêtes d'arrière plan :

Basé sur les théories de l'attention des premiers travaux sur la saillance visuelle [L. Itti et al], nous utilisons les nœuds sur la limite d'image comme les tiges d'arrière plan, c'est-à-dire, les données étiquetée (échantillons de requête) pour classer les pertinences de toutes les autres régions. Plus précisément, nous construisons quatre cartes de saillance utilisant les limites antérieures et les intégrons ensuite pour la carte finale, ce qui est désigné comme l'approche de Séparation / Combinaison (SC) : prendre la limite supérieure de l'image comme un exemple, nous utilisons les nœuds sur le côté comme des requêtes et les autres nœuds comme des données non étiquetées. Ainsi, on donne le vecteur d'indicateur \mathbf{y} , et tous les nœuds sont classés en fonction de l'équ.3 dans $\bar{\mathbf{f}}^*$, qui est un vecteur à N -dimension (N est le nombre total des nœuds du graphe). Chaque élément de ce vecteur indique la pertinence d'un nœud aux requêtes d'arrière-plan, et son complément est la mesure de saillance.

Nous normalisons ce vecteur à la gamme entre 0 et 1, et la carte de saillance en utilisant la limite antérieur supérieure, S_t peut être écrit comme:

$$S_t(i) = 1 - \bar{\mathbf{f}}^*(i) \quad i = 1, 2, \dots, N, \quad (5)$$

Où i indexe un nœud de superpixel sur le graphe, et $\bar{\mathbf{f}}^*$ désigne le vecteur normalisé.

De même, nous calculons les trois autres cartes S_b , S_l et S_r , en utilisant les limites du bas, de gauche et de droite d'image à comme des requêtes comme le montre la **fig.4**. Nous notons que les cartes de saillance sont calculées avec des vecteurs indicateurs différents \mathbf{y} tandis que le poids matrice \mathbf{W} et la matrice de degré \mathbf{D} sont fixés. Autrement dit, nous avons besoin de calculer l'inverse de la matrice $(\mathbf{D} - \alpha\mathbf{W})$ une seule fois pour chaque image. Puisque le nombre de superpixels est petit, la matrice inverse dans l'équ.3 peut être calculée efficacement. Ainsi, la charge de calcul global pour les quatre cartes est faible. Les quatre cartes de saillance sont intégrées par le processus suivant:

$$S_{bq}(i) = S_t(i) \times S_b(i) \times S_l(i) \times S_r(i). \quad (6)$$

II-4-2 Classement avec des requêtes d'avant plan :

La carte de saillance de la première étape est segmentée en binaire (c'est-à-dire l'avant et l'arrière plan sont saillants) en utilisant un seuil adaptatif, ce qui facilite la sélection des nœuds de l'avant plan des objets saillants comme des requêtes. Nous nous attendons à ce que les requêtes sélectionnées couvrent les régions d'objet saillantes, autant que possible (c'est-à-dire., avec le rappel élevé). Ainsi, le seuil est défini comme la saillance moyenne sur la carte de saillance entière. Une fois les requêtes saillantes sont données, un vecteur indicateur y est formé pour calculer le vecteur de classement f^* en utilisant l'équation 3. Comme cela est effectué dans le premier étape, le classement de vecteur f^* est normalisé entre la plage de 0 et 1 pour former la carte finale de saillance par :

$$S_{fq}(i) = \bar{f}^*(i) \quad i = 1, 2, \dots, N, \quad (7)$$

Où i indexe le nœud superpixel sur le graphe, et f désigne le vecteur normalisé.

Dans cette étape, nous notons qu'il y a des cas où les nœuds peuvent être mal choisis comme des requêtes de l'avant plan.

Tous ce traitement qu'on est entrain de faire avec un nombre de super pixel égal \emptyset on le refait pour un autre nombre qui égal a μ super pixels pour plus de précision sur l'objet saillant. et en fusionnant le résultat la carte de saillance résultante de 150 et de 1000 super pixels on génère une autre carte de saillance finale qui donne un résultat plus saillant et ceux avec l'équation suivante :

$$\text{Carte finale} = \gamma * \text{carte}_2 + \beta * \text{carte}_1 \quad (8)$$

Avec $\gamma = 0,3$, $\beta = 0,7$

Carte₂ : carte résultante de μ super pixels ; $\mu = 1000$

Carte₁ : carte résultante de \emptyset super pixels. $\emptyset = 150$

Le choix de ces paramètres a été pris de manière empirique.

Et pour résumer toutes les étapes du traitement on présente dans ce qui suit l'algorithme général qu'on a suivi.

II-5 Algorithme : la saillance sur la méthode multi-graphes basée sur Manifold Ranking

Entrée: Une image et les paramètres nécessaires

1: Segmenter l'image d'entrée en \emptyset superpixels, construire un graphe G avec des superpixels comme des nœuds, et calculer sa matrice de degré D et sa matrice de poids W par *l'équ.4*.

2: Calculer $(D - \alpha W)^{-1}$ et définir ses éléments diagonaux à 0.

3: Former des vecteurs indicateurs y avec des nœuds de chaque côté de l'image comme des requêtes, et calculer leurs cartes spécifiques secondaires correspondantes par *équ.3* et *équ.5*. Ensuite, calculer la carte de saillance S_{bq} par *l'équ.6*.

4: Bi-segmenter S_{bq} pour former des requêtes de l'avant plan saillants et un vecteur indicateur y . Calculer la carte de saillance S_{fq} par *l'équ.3* et *l'équ.7*.

Sortie1: une carte de saillance S_{fq1} représentant la valeur de saillance de chaque superpixel.

5 : Segmenter l'image d'entrée en μ superpixels, construire un graphe G avec des superpixels comme des nœuds, et calculer sa matrice de degré D et sa matrice de poids W par *l'équ.4*.

6: Calculer $(D - \alpha W)^{-1}$ et définir ses éléments diagonaux à 0.

7: Former des vecteurs indicateurs y avec des nœuds de chaque côté de l'image comme des requêtes, et calculer leurs cartes spécifiques secondaires correspondantes par *équ.3* et *équ.5*. Ensuite, calculer la carte de saillance S_{bq} par *l'équ.6*.

8: Bi-segmenter S_{bq} pour former des requêtes de l'avant plan saillants et un vecteur indicateur y . Calculer la carte de saillance S_{fq} par *l'équ. 3* et *l'équ.7*.

Sortie2: une carte de saillance S_{fq2} représentant la valeur de saillance de chaque superpixel.

9 : Fusionner la sortie1+sortie 2 avec l'équ.8

Sortie : une carte de saillance finale représentant ainsi l'objet saillant avec une meilleure précision.

Fin

II-6 Exemple d'application de l'algorithme









1. Input image 			
Segmentation avec 150 super pixels 		4- Segmentation avec 1000 super pixels 	
1.Carte de saillance pour le étape1	3-2 Carte de saillance pour le étape2	4-1 Carte de saillance pour le étape3	4-2 Carte de saillance pour le étape4
			
5- Fusion du résultat de 3-2 et 4-2 			

Fig6. Illustration des différentes étapes de notre algorithme

II-7 Résultats expérimentaux:

Nous comparons notre méthode avec quinze algorithmes de détection de saillance de l'état de l'art : Context and shape prior(CS)[JIANG et al.11], Human

visual sensitivity and amplitude spectrum (**VSAS**) [Yuming Fang et al.12], Segmentation driven low-rank matrix recovery (**SLR**) [Wenbin Zou et al.13], Hierarchical saliency detection (**HSL**) [Q. Yan et al.13], Joint spatial-color constraint and multi-scale segmentation (**SCC**) [Linfeng Xu et al.13], Saliency Tree (**ST**) [Zhi Liu et al.14], Unified approach to salient object detection via Low rank matrix recovery (**LMR**) [S.Xiaohui et W.Ying.12], Bayesian Saliency via Low and Mid Level Cues (**BS**) [Yulin Xie et al.11], Graph-Based Manifold Ranking (**MR**) [Bin Xu], Highlighting Sparse Salient Regions(**SSR**) [Xiaodi Hou et al.12], Principal component analysis (**PCA**) [M.Mudrovà, A.Procházka.11], Combining simple priors method (**CSP**) [Lin Zhang et al.13], Low-level features based on wavelet transform (**WT**) [N.Imamoglu, et al.13], Multisize superpixel approach for salient object detection based on multivariate normal distribution estimation (**MSA**) [Zhi Liu et al.14], Salient region detection via simple local and global contrast representation (**LGC**) [Jie Liu et al.14], Salient Region Detection with Soft Image Abstraction (**SIA**) [M.M. Cheng et al.13].

Dispositif expérimental:

Dans ce travail, nous utilisons la base d'images MSARA 1000 [R. Achanta. 11] (Microsoft public data base) qui contient des images couleurs, avec des luminances variantes, l'ombrage, avec leurs masques binaires, ces images sont prises grâce à des appareils différents, nous générons 1000 cartes de saillances avec une résolution complète pour chaque image d'entrée.

Nous avons fixé le nombre de nœuds de superpixel $\phi = 150$ puis $\mu = 1000$ nœuds dans toutes les expériences. Il des paramètres dans l'algorithme proposé: les poids de bord ω , θ en *équ.4*, et le poids d'équilibrage α dans *l'équ.3*. Les paramètres σ , ω , θ commandent la force de poids ; colorimétrique, variance et entropie entre une paire de nœuds : dans l'ordre et le paramètre α équilibre les contraintes de lissage et de montage dans la fonction de l'algorithme de régularisation de classement collecteur. Ces paramètres sont choisis de manière empirique $\sigma^2 = 0.1$, $\omega^2 =$, $\theta^2 =$, $\alpha = 0.99$, pour toutes les expériences.

Évaluation Métrique: Nous évaluons toutes les méthodes par la précision, le rappel et F-mesure. La valeur de précision correspond au rapport de pixels saillants correctement assignés à tous les pixels de régions extraites, tandis que la valeur de rappel est définie comme le pourcentage de pixels saillants détectées en ce qui concerne le nombre masque binaire.

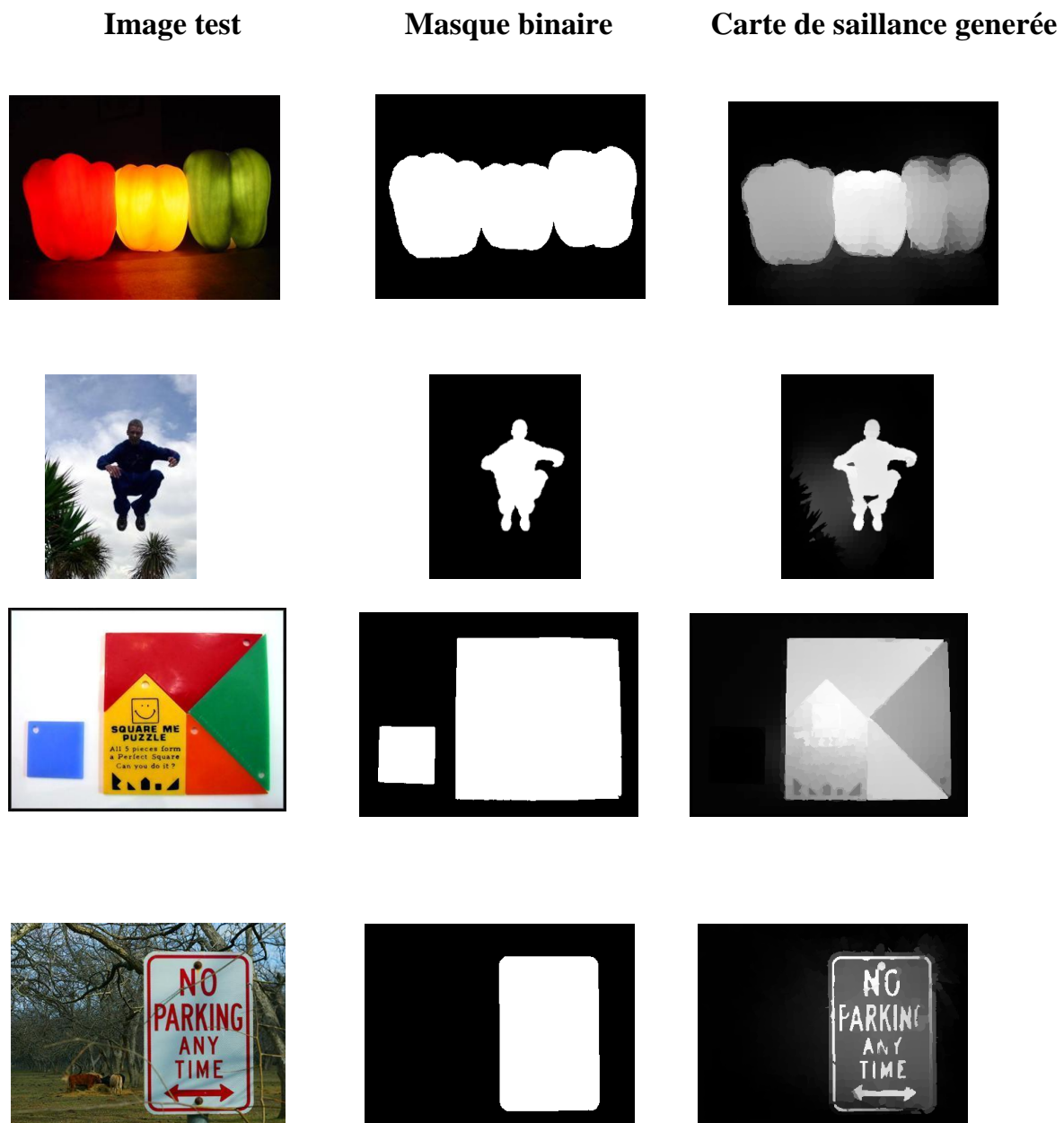


Fig7. Quelques résultats de notre travail

Semblable aux travaux préalables, les courbes de rappel de précision sont obtenues par la binarisation de la carte de saillance utilisant des seuils dans l'intervalle de 0 et 255. Le F-mesure est la mesure de la performance globale calculée par l'harmonique pondérée de précision et de rappel:

Pour la précision : $Précision = \frac{\sum_x g_x \times s_x}{\sum_x s_x}$

Pour le rappel : $Rappel = \frac{\sum_x g_x \times s_x}{\sum_x g_x}$

f-mesure $F_\beta = \frac{(1+\beta^2)Précision \times Rappel}{\beta^2 Précision + Rappel}, \quad (9)$

Avec : g : le masque binaire

s : la carte segmenter

x : pixel

Où nous mettons $\beta^2 = 0,3$ pour souligner la précision.

Nous calculons les courbes de précision rappel pour tous les algorithmes. Nous évaluons la performance de la méthode proposée contre quinze méthodes de l'état de l'art de détection de saillance. **La fig.8** montre les courbes de rappel précision de toutes les 15 méthodes citées au part avant. Nous notons que les méthodes proposées surpassent le ST [Zhi Liu et all.14], MSA [Zhi Liu et all.14], et HSL [Q. Yan et all.13] qui sont des méthodes de haut rendement pour la détection de saillance. Nous calculons aussi la précision, le rappel et F-mesure avec un seuil adaptatif, défini comme étant le double de la saillance moyenne de l'image. Le terrain le plus à droite de la **fig.8** montre que l'algorithme proposé atteint les valeurs de précision et F-mesure les plus élevées. Globalement, les résultats à l'aide de trois indicateurs montrent que l'algorithme proposé surpasse les méthodes de l'état de l'art.

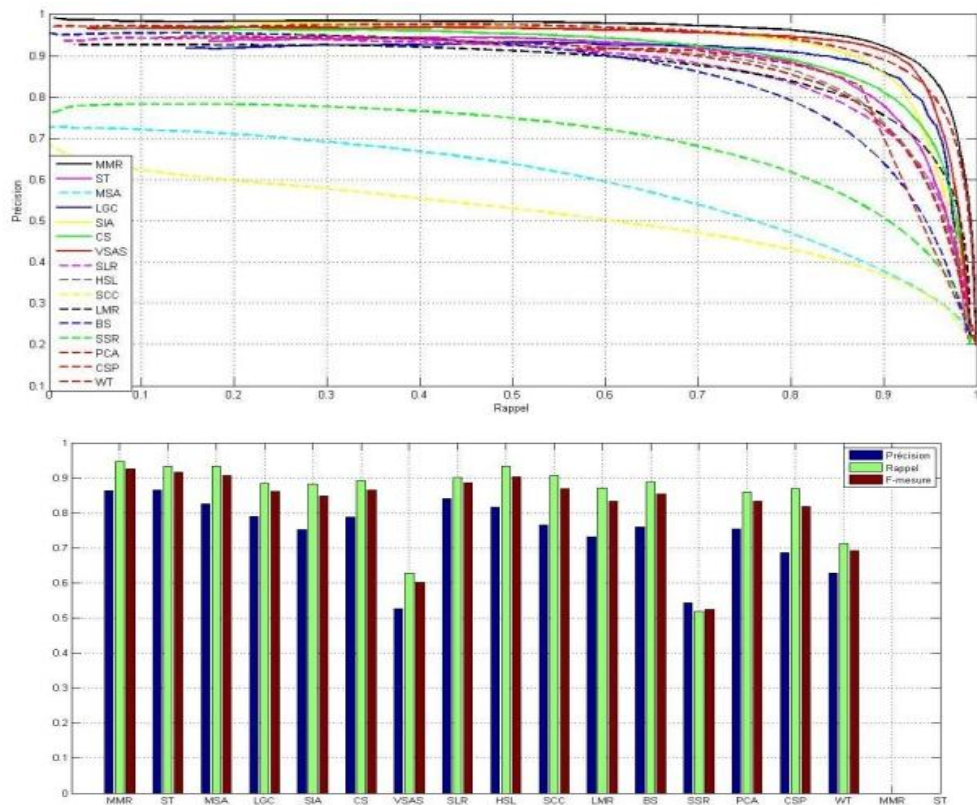


Fig.8 courbes précision rappel & la précision, le rappel et F-mesure : Tous les résultats sont calculés sur l'ensemble de données MSRA
Haut: courbes précision rappel des différentes méthodes de l'état de l'art avec la notre.
Bas: la précision, le rappel et F-mesure par seuil adaptatif

CONCLUSION :

Comme notre travail s'est basé sur le modèle développé par [Chuan Yang et al] qui est lui-même basé sur la méthode de multi-graphes (Manifold Ranking) qui intègre les indices de groupement locaux et les limites antérieurs de chaque pixel. Nous avons apporté quelques ajouts à la méthode pour avoir de meilleurs résultats qui consistent :

1. La transformation de l'espace de couleur du RGB vers le LAB pour plus de luminosité.
2. On a tenu compte de la texture, de la variance et de l'entropie de chaque nœud du super pixel.

3. On a considéré aussi la distance spatiale du centroïde entre deux supers pixels voisins.

4. Ainsi la segmentation à été faite pour deux seuils ($\emptyset \ll \mu$) super pixels car si on fait une segmentation avec \emptyset il y aura un manque d'information c.à.d. l'objet ne saura pas complètement saillant par contre si on applique une segmentation avec μ il y aura des résultats faux positifs, c'est pour cela qu'on a jugé intéressant de fusionner les deux.

5. Tous les traitements ont été faits pour ces deux seuils ce qui nous a donné deux cartes de saillance que nous avons par la suite fusionné par *l'équ8*.

L'évaluation de notre méthode par rapport à un ensemble de méthodes de l'état de l'art a bien montré que notre méthode est bien plus efficace ce qui se voit bien dans *la fig.8*.

Conclusion Générale

CONCLUSION GENERALE:

Dans ce manuscrit, nous nous sommes intéressés à la détection d'objets saillants sur les images en couleurs et on a fait une comparaison avec d'autres méthodes de l'état de l'art de hauts niveaux.

Nous avons en premier lieu porté notre réflexion sur la représentation des notions de généralité sur le traitement et la segmentation d'image. Nous avons proposé une représentation de traitement d'images qui est une discipline de l'informatique et des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information. Le mode et les conditions d'acquisition et de numérisation des images traitées conditionnent largement les opérations qu'il faudra réaliser pour extraire de l'information. En effet, de nombreux paramètres entrent en compte, les principaux étant : la résolution, les réglages optiques utilisés, qui déterminent par exemple la netteté de l'image, les conditions d'éclairage, qui déterminent une partie de la variabilité des images traitées, et le bruit de la chaîne de transmission d'image.

Nous avons ensuite proposé la segmentation d'image qui est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis. Les pixels sont ainsi regroupés en régions, qui constituent une partition de l'image. Il peut s'agir par exemple de séparer les objets du fond. Si le nombre de classes est égal à deux, elle est appelée aussi binarisation. Si l'homme sait naturellement séparer des objets dans une image c'est grâce à des connaissances de haut niveau (compréhension des objets et de la scène). Mettre au point des algorithmes de segmentation de haut niveau est encore un des thèmes de recherche les plus courants en traitement d'images.

En deuxième lieu on a porté notre réflexion sur la saillance qui est la manière de modéliser l'attention visuelle en informatique qui est elle-même la capacité naturelle de se concentrer de manière sélective sur un stimulus extérieur et de ne pas prendre en compte les informations voisines moins importantes. On peut distinguer deux approches, la première est dénommée « bottom-up » et la seconde « top-down » que nous avons très bien expliqué.

Conclusion Générale

En dernier lieu notre réflexion s'est portée sur notre méthode de multi-graphe basée sur manifold ranking. Nous avons proposé une méthode bottom-up pour détecter régions saillantes en images en couleur, une méthode qui a montré son efficacité par rapport à d'autres méthodes de l'état de l'art. Par ailleurs, cette méthode peut être améliorée pour avoir de meilleurs résultats chose que nous projetons de faire dans le futur.

BIBLIOGRAPHIE :

[A. *Traisman 80*]: A. Treisman and G. Gelade, “A feature-integration theory of attention” *Cognit. Psychol.*, vol. 12, no. 1, pp. 97–136, 1980.

[A.*Forsyth et J.Ponce.02*]: D. A. Forsyth et J. Ponce, « vision par ordinateur: une approche moderne ». Prentice Hall 2002.

[*Alexandre Hamon*]: Alexandre Hamon : « Segmentation couleur de plans cadastraux de la ville de Paris au XIXème siècle ».

[B. *Bruce. 03*]: B. Bruce & E. Jernigan. “Evolutionary design of context-free attentional operators”. In proc. ICIIP’03, pages 0–3. Citeseer, 2003.

[B. *Cocula et C. Peyroutet 86*]: Cocula B., Peyroutet C. « *Sémantique de l’image. Pour une approche méthodique des messages visuels* », Paris, Delagrave. 1986

[B. *Scholkopf et all.01*]: B. Scholkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson. “Estimating the support of a high-dimensional distribution”. *Neural Computation*, 2001.

[B. *Tatler.07*]: B. Tatler. The central fixation bias in scene viewing: “Selecting an optimal viewing position independently of motor biases and image feature distributions”. *Journal of Vision*, 2007.

[*Bin Xu*]: Bin Xu, Jiajun Bu, Chun Chen, Deng Cai, Xiaofei He, Wei Liu, Jiebo Luo, «Efficient Manifold Ranking for Image Retrieval».

[C. *Mozer. 98*]: Michael C. Mozer & Mark Sitton. “Computational modeling of spatial attention”. *Attention*, pages 341–393, 1998.

[C. *Zitnick et C. Kang.07*]: C. L. Zitnick and S. B. Kang. « Stereo for image-based rendering using image over-segmentation » *International Journal of Computer Vision (IJCV)*, 75:49–65, October 2007.

[*Coquerez. Philipp.95*]: J.p.Coquerez, S.Philipp. « Analyse d’images, filtrage et segmentation ». Edition Masson, Paris .1995.

[D. *Martin et all.01*]: D. Martin, C. Fowlkes, D. Tal, and J. Malik. « A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics ». In *IEEE International Conference on Computer Vision (ICCV)*, July 2001.

[D. *Zhou et all.03*]: D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Scholkopf. “Learning with local and global consistency”. In *NIPS*, 2003.

[D. Zhou et al.04]: D. Zhou, J. Weston, A. Gretton, O. Bousquet, and B. Scholkopf. “Ranking on data manifolds”. In NIPS, 2004.

[E. Conklin et D. McDonald 82]: Conklin E. J., McDonald D. D. “Saliency: The Key to the Selection Problem in Natural Language Generation”, In *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics*, Toronto, pp. 129-135.1982.

[E. Reiter et R. Dale 97]: Reiter E., Dale R, “Building Applied Natural-Language Generation Systems”, *Journal of Natural-Language Engineering*, 3, pp. 57-87. 1997.

[F, Lecumberry et al.10]: F, Lecumberry, A, Pardo et G. Sapiro. « Classification d'objets simultanée et la segmentation avec ordre élevé de multiples modèles de forme ». IEEE Transactions on Image Processing. pp: 625-635, 2010.

[F. Hamker. 05]: F Hamker. “Modeling Attention: From Computational Neuroscience to Computer Vision”. In *Attention and Performance in Computational Vision*, pages 118–132. Springer Berlin / Heidelberg, January 2005.

[F. Liu 06]: F. Liu & M. Gleicher. Region enhanced scale-invariant saliency detection. In *Proceedings of IEEE ICME*, pages 1–4. Citeseer, 2006.

[F. Meyer 88]: Meyer F, *Skeletons in digital spaces. Image analysis and mathematical morphology, theoretical advances.* Serra. Academic press.1988.

[F. Orabona .08]: Francesco Orabona, Giorgio Metta & Giulio Sandini. “A Protoobject based visual attention model”. In Lucas Paletta, editeur, *Attention in Cognitive Systems. Theories and Systems from an Interdisciplinary Viewpoint (WAPCV)*, pages 198–215, Berlin, Heidelberg, 2008. Springer. 2008.

[F. Perazzi et al.12]: F. Perazzi, P. Krahenbuhl, Y. Pritch, and A. Hornung. “Saliency filters: Contrast based filtering for salient region detection”. In *CVPR*, 2012.

[F.Boudjemai.09]: F.Boudjemai. « Comparaison des différents opérations de détection de contours ». Mémoire d'ingénieur en informatique, université Mouloud Mammeri de Tizi-Ouzou, 2009.

[G. Rousselet et M. Fabre-Thorpe 03]: .Rousselet G. A., Fabre-Thorpe M. « Les mécanismes de l'attention visuelle », *Psychologie Française*, 48, 1, pp. 29-44. 2003

[G. Sela et D. Marlin 97]: Gal Sela & Martin D Levine. « Real-time attention for robotic vision”. *Real-Time Imaging*, vol. 3, pages 173–194, 1997.

[H.Cettach.05]: H.Cettach. « Plate forme multi-agents pour la segmentation d'images ». Mémoire d'ingénieur. université Mouloud Mammeri de Tizi-Ouzou, 2005.

[J. Feng et al 11]: J. Feng, Y. Wei, L. Tao, C. Zhang, and J. Sun, “Salient object detection by composition,” in ICCV, 2011, pp. 1028–1035. 2011.

[J. Harel 06]: J. Harel, C. Koch, and P. Perona. “Graph-based visual saliency”. In NIPS, 2006.

[J. Itten 85]: Itten J « *Art de la couleur* », Paris, Dessain et Tolra.1985

[J. Russ.98]: J.C. Russ « The Image Processing Handbook», Third Edition, 1998.

[J. Sun et al.12] J. Sun, H. C. Lu, and S. F. Li. “Saliency detection based on integration of boundary and soft-segmentation”. In ICIP, 2012.

[J. Tsotsos 05]: J.K. Tsotsos, Y Liu, J Martineztrujillo, M Pomplun, E Simine & K Zhou. “Attending to visual motion. Computer Vision and Image Understanding”, vol. 100, no. 1-2, pages 3–40, 2005.

[J. McQueen.67]: J. McQueen. “Some methods for classification and analysis of multivariate observations”. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1:281-297, 1967.

[JIANG et al.11]: JIANG, WANG, YUAN, LIU, ZHENG, LI. «automatic salient object segmentation based on Context and Shape Prior, 2011».

[Jie Liu.14]: Jie Liu, Shengjin Wang. «Salient region detection via simple local and global contrast representation». doi:10.1016/j.neucom.2014.06.041.

[K. Kessler et al 96]: Kessler K., Duwe I., Strohner H. (1996) Sprachliche Objektidentifikation in ambigen Situationen : Empirische Befunde, *SFB 360 Situierete künstliche Kommunikation*, Report 96/1, Universität Bielefeld.

[K. Rapantzikos 03]: K Rapantzikos & N Tsapatsoulis. « On the implementation of visual attention architectures”. In *Tales of the Disappearing Computer* Santorini, 2003.

[K. Sayood. 00] : Khalid Sayood. Predictive coding, chapitre 6, pages 139–177. Morgan Kaufmann Publishers, 2nd édition, 2000.

[Kaddour, S. Aissa Brahim] C. Kaddour, S. Aissa Brahim, « Compression des images fixes par fractales basée sur la triangulation de delaunay et la quantification vectorielle ». Université des sciences et de la technique Houari Boumedine. Disponible sur www.kaddour.com.

[L. Itti 00]: Laurent Itti. “Models of Bottom-Up and Top-Down Visual Attention”. Phd, California Institute of Technology, 2000.

[L. Itti et al.98] : L. Itti, C. Koch, and E. Niebur. “A model of saliency-based visual attention for rapid scene analysis”. IEEE PAMI, 1998.

[L. Itti.04] : L. Itti. “Automatic foveation for video compression using a neurobiological model of visual attention”. IEEE TIP, 2004.

[L. Ladicky et al.09]: L. Ladicky, C. Russell, P. Kohli, P. Torr. « Associatifs ERC hiérarchiques pour objet Classe image Segmentation». Dans: ICCV 2009.

[L. Marchesotti 10]: L. Marchesotti, C. Cifarelli, and G. Csurka, “A framework for visual saliency detection with applications to image thumbnailing,” in Proceedings of IEEE International Conference on Computer Vision, 2010, pp. 2232–2239.2010

[L. Vincent 91]: Vincent L, Soille P, Watershed in digital spaces, an efficient algorithm based on immersion simulation. Trans. PAMI vol 13, n° 6, jun 91.

[L. Zhang et al 08]: L. Zhang, M. Tong, T. Marks, H. Shan, and G. Cottrell, “Sun: “A Bayesian framework for saliency using natural statistics”, Journal of Vision, vol. 8, no. 7, 2008.

[L.Grady et al.11] : L. Grady, M. Jolly, and A. Seitz. “Segmenation from a box”. In ICCV,2011.

[L.Wang et al.11] : L. Wang, J. Xue, N. Zheng, and G. Hua. “Automatic salient object extraction with contextual cue”. In ICCV, 2011.

[L.Zhu et al.14]: Zhu L, Klein DA, Frintrop S, Cao Z, Cremers AB. «A multisize superpixel approach for salient object detection based on multivariate normal distribution estimation». IEEE Trans Image Process. 2014 Dec;23(12):5094-107. doi: 10.1109/TIP.2014.2361024. Epub 2014 Oct 1.

[Lin Zhang et al.13]: Lin Zhang, Zhongyi Gu, and Hongyu Li,SDSP. «A novel saliency detection method by combining simple priors». School of Software Engineering, Tongji University, Shanghai, China, 2013 20th IEEE International Conference on Image Processing (ICIP).

[Linfeng Xu et al.13]: Linfeng Xu, Hongliang Li, Liaoyuan Zeng, King Ngi Ngan, «Saliency detection using joint spatial-color constraint and multi-scale segmentation» Published in: Journal of Visual Communication and Image Representation, Volume 24 Issue 4, May, 2013, Pages 465-476.

[Ludovic, Sylvie] : Ludovic Macaire, Sylvie Philipp-Foliguet « Segmentation d’images couleur ».

[M. Aziz 09]: M. Aziz & B. Mertsching. “Early Clustering Approach towards Modeling of Bottom-Up Visual Attention”. In KI 2009: Advances in Artificial Intelligence, volume

1289, pages 315–322. Springer, 2009.

[M. Everingham et al.10]: M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. « The PASCAL Visual Object Classes Challenge». International Journal of Computer Vision (IJCV), 88(2):303–338, June 2010.

[M. Freeman 89]: Freeman .M “*L’Image*”, Paris, Editions VM. 1989.

[M. Lopez 06] : M Lopez, A Fernandezcaballero, M Fernandez, J Mira & A Delgado. « Motion features to enhance scene segmentation in active visual attention”. Pattern Recognition Letters, vol. 27, no. 5, pages 469–478, 2006.

[M. Mancas 07]: Matei Mancas. « Computational Attention : Towards attentive computers ». Phd, Faculté Polytechnique de Mons, 2007.

[M. Ran 13]: Ran Margolin, Ayellet Tal, Lihi Zelnik-Manor, " What Makes a Patch Distinct? ", CVPR 2013.

[M. Segall et al 66] :Segall M. H., Campbell D. T., Herskovits M. J. “*The Influence of Culture on Visual Perception*”, Indianapolis, Bobbs-Merrill. 1966

[M. Sonka et al.08]: M. Sonka, V. Hlavac et R. Boyle. «Image Processing, Analysis, and Machine Vision» Thomson. 2008.

[M. Vettraino-Soulard 93] : Vettraino-Soulard M.-C. « *Lire une image* », Paris, Armand Colin. 1993

[M.M. Cheng et al. 13]: Ming-Ming Cheng, Jonathan Warrell , Wen-Yan Lin, Shuai Zheng, Vibhav Vineet Nigel Crook, Vision Group. «Efficient Salient Region Detection with Soft Image Abstraction». Oxford Brookes University.2013

[M.Mudrovà, A.Procházka.11]: M.Mudrovà, A.Procházka. «Principal component analysis in image processing». Institute of chemical technology, prague department of computing and control engineering, 2011.

[M.Ouali. K.Mohamed.09] : M.Ouali, K.Mohamed. « Segmentation d’image, multi-composantes par coopération des méthodes ». Mémoire d’ingénieur, en informatique, université Mouloud Mammeri de Tizi-Ouzou, 2009.

[MM Cheng et al.11] :] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. M. Hu.”Global contrast based salient region detection”. In CVPR, 2011.

[N. Otsu. 79]: N. Otsu *A threshold selection method from grey scale histogram*, IEEE Trans. on Syst. Man and Cyber., vol 1, pp 62-66, 1979

[N.Imamoglu, N et all. 13] : Imamoglu, N., Weisi Lin ; Yuming Fang. «A Saliency Detection Model Using Low-Level Features Based on Wavelet Transform», Published in: Multimedia, IEEE Transactions on (Volume:15 , Issue: 1),pages 96 -105.2013

[O. Le Meur 05] : O. Le Meur, D. Thoreau, P. Le Callet & D. Barba. « A spatio-temporal model of the selective human visual attention”. In Image Processing, 2005. ICIP 2005. IEEE International Conference on, volume 3, pages III–1188. IEEE, 2005.

[P. Edmonds 93]: Edmonds P. G. A “Computational Model of Collaboration on Reference in Direction-Giving Dialogues”, Ms. Thesis, University of Toronto. 1993

[P. J. B. Hangcock et all.92]: P. J. B. Hangcock, R. J.Baddely, and L. S Smith. «The principal components of natural images». Network, 3:61 – 70, 1992.

[P. Matthieu 10 :] Matthieu Perreira Da Silva, « Modèle computationnel d’attention pour la vision adaptative », P.H.D thes is in computer science and application, Rochelle university, 2010.

[Q. Yan et all.13]: Q. Yan, L. Xu, J. Shi, and J. Jia. “Hierarchical saliency detection”. In CVPR, 2013.

[R. Achanta et all.08] : R. Achanta, F. Estrada, P. Wils, and S. “Susstrunk. Salient region detection and segmentation”. In ICVS, 2008.

[R. Achanta et all.09] : R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk. “Frequencytuned salient region detection”. In CVPR, 2009.

[R. Achanta et all.10] : R. Achanta, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. “Slic superpixels”. Technical report, EPFL, Tech.Rep. 149300, 2010.

[R. Achanta.11]: R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, «SLIC Superpixels Compared to State-of-the-art Superpixel Methods». journal of latex class files, vol. 6, no. 1, December 2011.

[R. Desimone 95]: R Desimone & J Duncan. “Neural mechanisms of selective visual attention”. Annual review of neuroscience, vol. 18, pages 193–222, January 1995.

[R. Fisher.25]:Ronald Fisher, « *The influence of rainfall on the yield of wheat at Rothamsted* », Philosophical Transactions of the Royal Society, 1925.

[R. Milanese 94]: R. Milanese, H. Wechsler, S. Gill, J.-M. Bost & T. Pun. “Integration of bottom-up and top-down cues for visual attention using non-linear relaxation”. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pages

781–785. IEEE Comput. Soc. Press, 1994.

[**S. Ahmad 92**]: Subutai Ahmad. “VISIT: An efficient computational model of human visual”. 1992

[**S. Choi 06**]: Sang-bok Choi, Bum-soo Jung & Sang-woo Ban. “Biologically motivated vergence control system using human-like selective attention model”. *Neurocomputing*, vol. 69, no. 4-6, pages 537–558, January 2006.

[**S. Dong 06**]: Dong, S.W. Ban & Minho Lee. “Biologically Inspired Selective Attention Model Using Human Interest”. *International Journal of Information Technology*, vol. 12, no. 2, pages 140–148, 2006.

[**S. Gilles 96**]: Sebastien Gilles. “Description and experimentation of image matching using mutual information”, 1996.

[**S. Goferman et al.10**] : S. Goferman, L. Zelnik-Manor, and A. Tal. Context aware saliency detection. In *CVPR*, 2010.

[**S.Xiaohui and W.Ying.12**]: Xiaohui Shen and Ying Wu, «A Unified Approach to Salient Object Detection via Low Rank Matrix Recovery». In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[**T. Chen et al.09**] : T. Chen, M. Cheng, P. Tan, A. Shamir, and S. Hu. “Sketch 2 photo”: Internet image montage. *ACM Trans. on Graphics*, 2009.

[**T. Geerinck 09**]: Thomas Geerinck, Hichem Sahli, David Henderickx, Iris Vanhamel & Valentin Enescu. “Modeling Attention and Perceptual Grouping to Salient Objects”. In *Attention in Cognitive Systems*, page 166. Springer London, Limited, 2009.

[**T. Kadir. 01**]:Timor Kadir & Michael Brady. “Saliency, scale and image description”. *International Journal of Computer Vision*, vol. 45, no. 2, pages 83–105, 2001.

[**T. Pattabhiraman 93**] : Pattabhiraman T. “*Aspects of Saliency in Natural Language Generation*”, Ph.D. Thesis, Simon Fraser University. 1993

[**T.Liu et al.11**] : T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum. «Learning to detect a salient object». *IEEE PAMI*, 2011.

[**U. Rutishauser et al.04**] : U. Rutishauser, D. Walther, C. Koch, and P. Perona. Is bottom-up attention useful for object recognition in *CVPR*». 2004.

[**V. Gopalakrishnan, et al.10**]: V. Gopalakrishnan, Y. Hu, and D. Rajan. “Random walks on graphs for salient object detection in images”. *IEEE TIP*. 2010

[**V. Lempitsky et al.09**] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp. “Image

segmentation with a bounding box prior”. In ICCV, 2009.

[Wenbin Zou et al.13]: Wenbin Zou, Kidiyo Kpalma, Zhi Liu, Joseph Ronsin, «Segmentation Driven Low-rank Matrix Recovery for Saliency Detection». 2013

[Xiaodi Hou et al.12]: Xiaodi Hou, Jonathan Harel, and Christof Koch, Member, IEEE, «Highlighting Sparse Salient Regions», IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 1, january 2012.

[Xie et al.13]: Y.L. Xie, H.C. Lu, and M.H. Yang. « Bayesian saliency via low and mid level cues”. IEEE TIP, 2013.

[Y. Baticle 85]: Y.Baticle «Clés et codes de l’image »,Paris, Magnard Université. 1985

[Y. C. Wei et al.12]: Y. C. Wei, F. Wen, W. J. Zhu, and J. Sun. “Geodesic saliency using background priors”. In ECCV, 2012.

[Y. Chuan .13]: Chuan Yang, Lihe Zhang, Huchuan Lu, Xiang Ruan and Ming-Hsuan Yang, "Saliency Detection via Graph-Based Manifold Ranking", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2013), Portland, June, 2013.

[Y. Lu et al.11]: Y. Lu, W. Zhang, H. Lu, and X. Y. Xue. “Salient object detection using concavity context”. In ICCV, 2011.

[Yulin Xie et al.11]: Yulin.Xie, Huchuan.Lu, and Ming-Hsuan Yang. «Transactions on image processing».2011, IEEE

[Yuming Fang et al.12]: Yuming Fang, Weisi Lin, Bu-Sung Lee, Chiew Tong Lau, Zhenzhong Chen, Chia-Wen Lin. «Bottom-Up Saliency Detection Model Based on Human Visual Sensitivity and Amplitude Spectrum». IEEE Transactions on Multimedia 14(1): 187-198. 2012.

[Z. Pylyshyn. 01]: Pylyshyn Z. W. “Visual Indexes, Preconceptual Objects, and Situated Vision”, *Cognition*, 80, 1/2, pp. 127-158. 2001.

[Zhi Liu et al.14]: Zhi Liu, Member, IEEE, Wenbin Zou, and Olivier Le Meur, «Saliency Tree: A Novel Saliency Detection Framewor», IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 23, NO. 5, MAY 2014.