



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique

Université Mohamed Khider – BISKRA

Faculté des Sciences Exactes, des Sciences de la Nature et de la Vie

**Département d'informatique**

N° d'ordre : SIOD 7/M2/2019

**Mémoire**

Présenté pour obtenir le diplôme de master académique en

**Informatique**

Parcours :  **Systèmes d'information, Décision et Optimisation**

---

**Détection de l'attention visuelle  
par la saillance basée sur la  
théorie de décision**

---

Par :

**ROUINA AMINA**

Soutenu le 07 juillet 2019, devant le jury composé de :

DJEDI Noureddine

Professeur

Président

BABAHENINI Mohamed Chaouki

Professeur

Rapporteur

TORKI Fatima Zohra

MAA

Examineur

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

# Dédicaces

*Avant tout, je remercie sincèrement Allah, le Tout-Puissant, de m'avoir donné la capacité et la patience suffisantes pour mener à bien ce mémoire.*

*Je dédie ce travail à mes chers parents qui m'ont beaucoup soutenu, physiquement et moralement.*

*À mes chers frères et sœurs : Med Lamine, Messaoud, Chouaib, Oussama, Rokaia et Asma.*

*À tous mes amies : Saouab, Samah, Hanane et Amira*

*Et à tout ce que je connais sans exceptions.*

*A tous mes professeurs*

*À toute ma famille élargie ROUINA*

*Enfin, j'adresse mes salutations et mes bénédictions à tous ceux qui m'ont soutenu dans l'accomplissement de ce travail.*

# Remerciements

La première et dernière chose est pour Allah qui me fournit la capacité suffisante pour terminer ce travail.

J'aimerais remercier mon encadreur, le Professeur **BABAHENINI Mohamed Chaouki** pour les conseils, les encouragements prodigués tout au long de mon parcours en tant qu'étudiante. J'ai eu beaucoup de chance d'avoir un superviseur qui se souciait tellement de mon travail et qui a répondu à mes questions et requêtes si rapidement.

Je tiens à remercier aussi **Dr SAOULI Abdelhak** qui m'a beaucoup aidé pour réaliser ce travail.

Je n'oublierai pas de remercier les membres du jury Pr **DJEDI Noureddine** et Madame **TORKI Fatima Zohra** et Pr **CHERIF Foudil** qui ont accepté de lire et d'examiner ce travail.

Je suis infiniment reconnaissante envers mes enseignants, qui m'ont fait profiter de leur savoir, qui m'ont encouragée et aidée durant mon cursus universitaire.

Un merci spécial à tous ceux qui m'ont soutenu pour terminer ce travail.



## Résumé

La perception visuelle est l'impression captée par les yeux, qui modélise les informations les plus pertinentes dans une image quelconque. Son but est de choisir quelle zone de l'image va être analysée avant les autres car elle serait potentiellement plus intéressante. Ainsi elle permet de réduire la quantité d'informations à traiter, et par conséquent d'accélérer l'ensemble du processus de vision.

Dans ce mémoire, nous avons proposé d'étudier et d'implémenter un modèle de perception « Modèle de théorie de décision » afin de produire une carte (image 2D à niveau de gris) de saillance qui sera utilisée dans les traitements postérieures pour l'amélioration des images, ce qui va permettre de faire des traitements selon l'importance de la perception visuelle représentation conduisant à trouver le bon compromis entre la bonne qualité des traitements et le temps de calcul de l'image résultat.

## Abstract

Visual perception is the impression captured by the eyes, which models the most relevant information in any image. Its purpose is to choose which area of the image will be analyzed before the others because it would be potentially more interesting. Thus it reduces the amount of information to be processed, and therefore accelerates the entire process of vision.

In this thesis, we proposed to study and implement a "decision theory model" perception model to produce a salient greyscale 2D image that will be used in posterior processing. image enhancement, which will allow to make treatments according to the importance of the visual perception representation leading to find the right compromise between the good quality of the treatments and the calculation time of the result image.

## الملخص

الإدراك البصري هو الانطباع الذي تم التقاطه بواسطة العيون ، والذي يصمم أكثر المعلومات صلة بأي صورة. والغرض منه هو اختيار أي جزء من الصورة سيتم تحليله قبل الآخرين لأنه سيكون أكثر إثارة للاهتمام. وبالتالي فإنه يقلل من كمية المعلومات التي يتعين معالجتها ، وبالتالي يسرع عملية الرؤية بأكملها

في هذه الرسالة ، اقترحنا دراسة وتنفيذ نموذج إدراك "نموذج نظرية القرار" لإنتاج صورة رمادية ثنائية الأبعاد بارزة سيتم استخدامها في المعالجة اللاحقة. تحسين الصورة ، والذي سيتيح إجراء المعالجات وفقاً لأهمية تمثيل الإدراك البصري مما يؤدي إلى إيجاد حل وسط مناسب بين الجودة الجيدة للمعالجات ووقت حساب الصورة الناتجة

## Liste des tables

**Table 01 :** Table de comparaison les résultats de modèle Itti avec le modèle de discriminante.....53

**Table 02 :** Table de comparaison les résultats de modèle Itti avec le modèle de discriminante (Base de données : images de la base d'entraînement) ..... 54

**Table 03 :** Table de comparaison les résultats de modèle Itti avec le modèle de discriminante (Base de données : images d'Art) .....55

**Table 04 :** Table de comparaison les résultats de modèle Itti avec le modèle de discriminante de nombre de caractéristique variable.....55

## Liste de Figures

<b>Figure 1</b> : Traitement de l'information visuelle dans le système visuel humain.....	5
<b>Figure 2</b> : (a) La structure de l'œil ; (b) La structure laminaire de la rétine qui montre les cinq premières couches. La lumière doit traverser ces couches pour atteindre les photorécepteurs situés à l'arrière de la rétine.....	6
<b>Figure 3</b> : Spectre des ondes électromagnétiques et différents domaines d'application (extrait du cours de Physique Générale PhyG II-7) .....	7
<b>Figure 4</b> : Couleurs Additives (Les couleurs primaires rvg).....	9
<b>Figure 5</b> : couleurs soustractives (couleurs primaire cmj « film, imprimante »).....	9
<b>Figure 6</b> : Système RVB.....	10
<b>Figure 7</b> : Diagramme de chromaticité CIE (x, y). La coordonnée trichromatique x est en abscisse et y en ordonnée.....	11
<b>Figure 8</b> : CIE XYZ (fonction d'observation d'un standard).....	11
<b>Figure 9</b> : TSL /TSV (Teinte Saturation Luminance / Teinte Saturation Valeur).....	12
<b>Figure 10</b> : Espace $l^* a^* b^*$ .....	13
<b>Figure 11</b> : Classification des modèles de l'attention visuelle.....	15
<b>Figure 12</b> : Approche ascendante (Bottom-up).....	19
<b>Figure 13</b> : Modèle de carte de saillance Koch &Ullman. (Extrait de [Koch et Ullman, 1987]).....	21
<b>Figure 14</b> : Le modèle décrit par Itti et al. (1998) puis par Itti & Koch (2001).....	22
<b>Figure 15</b> : Modèle Bayésien.....	23
<b>Figure 16</b> : Modèle de Théorie de Décision.....	23
<b>Figure 17</b> : Modèle de Théorie de l'information.....	24
<b>Figure 18</b> : Modèle d'analyse Spectrale.....	25
<b>Figure 19</b> : (a) (b) (c) Divers exemples difficiles pour les détecteurs de saillance actuels. a) pomme suspendue à un arbre, (b) un oiseau dans un arbre, (c) un œuf dans un nid. (d) certaines fonctions de base DCT. De gauche à droite, de haut en bas, es détecteurs	

de : bords, barres, coins, jonctions en T, points, plages de flux et clés à contrôle.....	29
<b>Figure 20</b> : Filtrage de l'image I par le filtre H.....	33
<b>Figure 21</b> : Filtre de convolution.....	34
<b>Figure 22</b> : Filtrage par convolution discrète 2D.....	35
<b>Figure 23</b> : Problème gaussien avec deux classes $Y = \{1, 2\}$ , dans les deux dimensions, $X = (X1, X2)$ .....	36
<b>Figure 24</b> : (a) L'image original, (b) Histogramme de l'image.....	37
<b>Figure 25</b> : Schéma globale de modèle de Discriminante.....	42
<b>Figure 26</b> : Logo de Visual Studio 2010.....	46
<b>Figure 27</b> : Diagramme de Classe de modèle implémenter.....	46
<b>Figure 28</b> : Capture d'écran présente notre application en mode de fonction qui affiche les images (a) image original, (b) carte de saillance Itti, (c) carte de saillance du modèle discriminante.....	47
<b>Figure 29</b> : Capture d'écran présente la procédure qui charge les caractéristiques et calcule la DCT.....	48
<b>Figure 30</b> : Capture d'écran présente la procédure de filtrage a convolution	
<b>Figure 31</b> : Capture d'écran présente la procédure de Calcule les Réponses ( $R_i$ ).....	49
<b>Figure 32</b> : Capture d'écran présente la procédure de Calcule MMD.....	50
<b>Figure 33</b> : Capture d'écran présente la fonction de calcule l'histogramme.....	51
<b>Figure 34</b> : Capture d'écran présente la procédure de calcule PDF.....	51
<b>Figure 35</b> : Capture d'écran présente la procédure de calcul de la saillance.....	52

## Liste d'Equations

(1) L'équation générale pour un DCT 1D.....	32
(2) L'équation générale pour un DCT 2D.....	32
(3) Théorème de Bayes naïf.....	37
(4) Probabilité postérieure.....	38
(5) La règle de décision pour Maximiser la probabilité postérieure.....	39
(6) Probabilité de classe conditionnelle.....	39
(7) Probabilité préalable .....	40
(8) Preuve.....	40
(9) Lissage Additif.....	41
(10) Calcule de la saillance (La somme) .....	41

## Table des matières

Remerciements

Dédicaces

Résumé

Absract

Liste de tables

Liste de Figures

Liste d'Equations

Table des matières

Introduction General ..... 1

### Chapitre 01 : Perception visuelle et Couleur

1. Introduction.....	4
2. La Neurologie Contexte de la Vision et de L'attention Anticipées.....	4
3. Système Visuel Humain (SVH).....	4
4. La Structure des Yeux.....	5
4.1. Composants de l'œil.....	5
4.1.1. La sclérotique.....	5
4.1.2. La choroïde.....	5
4.1.3. La rétine.....	6
5. La perception visuelle.....	7
6. La perception des couleurs.....	7
6.1. Couleur.....	7
7. Les systèmes de couleurs.....	7
7.1. Le système colorimétrique RVB (ou RGB en anglais).....	9
7.2. Le système CIE XYZ.....	10
7.3. Le système TSL /TSV (Teinte Saturation Luminance / Teinte Saturation Valeur).....	12
7.4. L'espace colorimétrique absolu CIE L*a*b 1976.....	12
8. Conclusion.....	13

### Chapitre 02 : Classification des modèles de l'attention visuelle

1. Introduction.....	15
2. Attention Visuelle.....	16
2.1. Psychophysique.....	16
2.2. Neurophysiologie.....	16
2.3. Modélisation Computationnelle.....	16

2.3.1. Modèles neuronaux.....	17
2.3.2. Modèles de Filtres.....	17
2.3.2.1. Modèles de Top-down.....	17
2.3.2.1.1. Modèles de recherche visuelle.....	17
2.3.2.1.2. Modèles de contextes (context (Gist) models) .....	18
2.3.2.1.3. Modèles axés sur les tâches(Task-driven models).....	19
2.3.2.2. Modèles de Bottom-up.....	19
2.3.2.2.1. Carte de Saillance.....	20
3. Modelés d'attention.....	21
3.1. Modèles Cognitifs.....	21
3.2. Modèles Bayésiens.....	22
3.3. Modèles Théorique de décision.....	23
3.4. Modèles de Théorie de l'information.....	24
3.5. Modèles Graphiques.....	24
3.6. Modèles d'Analyse Spectrale.....	25
3.7. Modèles de Classification.....	25
4. Bilan.....	26
5. Conclusion.....	27

### Chapitre 03 : Technique de l'attention Visuelle basée sur la théorie de décision.

1. Introduction.....	29
2. Détection de saillance.....	29
3. La saillance discriminante.....	30
3.1. Sélection des caractéristiques évolutives.....	31
3.2. Modèles biologiquement plausibles.....	31
3.2.1. Etape 01 : Extraction de caractéristiques (Calcul DCT).....	31
3.2.1.1. Transformation en cosinus discrète (DCT).....	33
3.2.1.2. Transformée de Fourier discrète(DFT).....	32
3.2.2. Etape 02 : Calculer $R_i$ .....	33
3.2.2.1. Filtrage.....	33
3.2.2.2. Filtrage par convolution discrète 2D.....	34
3.2.3. Etape 03 : Calculer les poids $w_i$ .....	35
3.2.3.1. Maximum marginal diversity.....	35
3.2.3.2. Histogramme.....	36
3.2.3.3. Théorème de Bayes naïf.....	37
3.2.3.3.1. Probabilité postérieure.....	38
3.2.3.3.2. Probabilité de classe conditionnelle.....	39
3.2.3.3.3. Probabilité préalable.....	39
3.2.3.3.4. Preuve.....	40
3.2.3.3.5. Lissage additif.....	41
3.2.4. Etape 04 : calcul de la saillance (la somme) .....	41
4. Schéma Global.....	41
5. Conclusion.....	42

### Chapitre 04 : Implémentation,résultats et bilan.

1. Introduction.....	44
2. Outil de développement.....	44

2.1. Langages de programmation.....	44
2.2. Bibliothèques utilisées.....	45
3. Implémentation.....	46
3.1. Diagramme de classe de modèle conceptuel.....	46
3.2. Procédures et Fonctions utilisées.....	48
4. Teste et Comparaison.....	52
4.1. Test de résultats Visuels.....	52
4.1.1. Base de teste : Images du Texture.....	53
4.1.2. Base de test : Images de la base d'entraînement.....	54
4.1.3. Base de test : Images d'Art.....	55
4.2. Test de Résultats par rapport le nombre de Caractéristiques.....	56
5. Conclusion.....	58

Conclusion General

Bibliographie et webographie

# Introduction Générale

## Contexte général du travail

La perception visuelle est considérée comme un outil de gestion de l'information ; elle permet d'effectuer la modélisation des parties attirantes selon la vision humaine. La détection de la saillance de l'image se concentre sur l'extraction des informations pertinentes de l'image, ainsi l'extraction des régions de saillance est un modèle de perception visuelle, nécessaire dans la plupart des méthodes de traitement d'image basées sur le contenu de l'image, car les composants importants de l'image fournissent l'information la plus complète sur une image entière. Par conséquent, l'extraction des régions saillantes des images facilite efficacement de nombreuses applications telles que la récupération d'image, la compression d'image adaptative, la reconnaissance d'objet, le redimensionnement d'image, par un traitement selon le niveau d'importance de l'attention visuelle.

Les humains peuvent facilement se concentrer sur les parties saillantes des images selon l'expérience et le jugement, mais les machines sont incapables de reproduire précisément une telle capacité. De nombreux chercheurs ont étudié cette question sur la base de la biologie, la physiologie et la neurobiologie. Dans ces études, certaines caractéristiques que les régions saillantes devraient avoir, y compris l'unicité, le caractère aléatoire, et des caractéristiques surprenantes, sont acquises.

## Problématique et objectifs

Notre objectif est donc de mettre en œuvre la technique de théorie de décision et ensuite de comparer les résultats obtenus avec ceux de la technique de base d'Itti et Koch.

Dans ce mémoire, nous avons proposé d'étudier et d'implémenter un modèle de perception « Modèle de théorie de décision » afin de produire une carte (image 2D à niveau de gris) de saillance qui sera utilisée dans les traitements postérieures pour l'amélioration des images, ce qui va permettre de faire des traitements selon l'importance de la perception visuelle représentation conduisant à trouver le bon compromis entre la bonne qualité des traitements et le temps de calcul de l'image résultat.

## Organisation du mémoire

Notre mémoire se compose de 04 chapitres :

Chapitre 01 : nous présentons Le Contexte Neurologique de la Vision et de L'attention Anticipées, puis nous parlons sur le système visuel humain et l'œil, ensuite nous présentons la perception visuelle et la perception de couleur, nous exposons en particulier la notion de couleur vue qu'elle présente un élément essentiel du système perceptuel.

Chapitre 02 : Nous allons présenter, une classification des modèles de l'attention visuelle, nous avons terminé par une synthèse dressant un bilan sur les caractéristiques, avantages et inconvénients des modèles.

Chapitre 03 : Dans ce chapitre nous allons détailler, les principaux outils utilisés dans la mise œuvre de la technique de théorie de décision, ainsi nous exposons les algorithmes utilisés pour le filtrage en utilisant les caractéristiques extraits à partir de la DCT et quelques rappels sur les probabilités conditionnelles

Chapitre 04 : Concerne la partie implémentation, dans laquelle nous allons présenter les langages de programmation et les outils de développement utilisés lors de l'implémentation de notre application, puis nous allons exposer le diagramme de classe de modèle conceptuel, procédures et fonctions utilisées dans la partie de l'implémentation. Pour la validation, nous comparons les résultats obtenus après la réalisation de notre application aux résultats obtenus par la méthode de base d'Itti. Une deuxième validation a été faite en changeant le nombre de caractéristiques pour voir la variation des résultats.

# **Chapitre 01**

## **Perception visuelle et Couleur**

## 1. Introduction

Nous nous sommes posé une question : "Le cerveau est-il vraiment influencé par notre perception visuelle des couleurs ?", par conséquent nous avons tenté d'y trouver une réponse. La perception visuelle est un procédé qui permet d'interpréter les informations qui caractérisent notre environnement. Ce mécanisme met en relation deux organes responsables du sens de la vue : l'œil et le cerveau.

Dans ce chapitre, nous présentons La Neurologie Contexte de la Vision et de L'attention Anticipées, puis nous parlons sur le système visuel humain et l'œil, ensuite nous présentons la perception visuelle et la perception de couleur, nous exposons en particulier la notion de couleur vue qu'elle présente un élément essentiel du système perceptuel.

## 2. Le Contexte Neurologique de la Vision et de L'attention Anticipées

Le système visuel est le plus complexe de tous les systèmes sensoriels et une grande partie du cerveau humain est impliquée dans la vision [1]. En raison de sa complexité, l'étude du système visuel a mobilisé les efforts d'un grand nombre de scientifiques et des progrès importants ont été réalisés au cours des dernières décennies [1]. L'étude du système visuel permet d'identifier les principes communs du traitement de l'information sensorielle. Les découvertes empiriques en neurobiologie et physiologie servent de motivation et de guide pour les modèles informatique.

## 3. Système Visuel Humain (SVH)

SVH est un système très complexe pas encore complètement compris. La vision est l'un de nos sens les plus délicats et les plus compliqués. Elle est également étudiée de manière intensive. Selon [2], un quart du cerveau est impliqué dans le traitement de l'information visuelle. On en sait plus sur la vision que tout autre système sensoriel chez les vertébrés et toutes ces informations sont issues d'études sur des singes et des chats. Certaines parties de base du SVH sont brièvement décrites dans les sections suivantes. Les principaux éléments de base qui jouent un rôle majeur dans le SVH sont les yeux, le noyau géniculé latéral (NGL), le cortex visuel (V1, V2, V3, V4), la région temporale médiane (MT) et le cortex inférotemporal (IT).

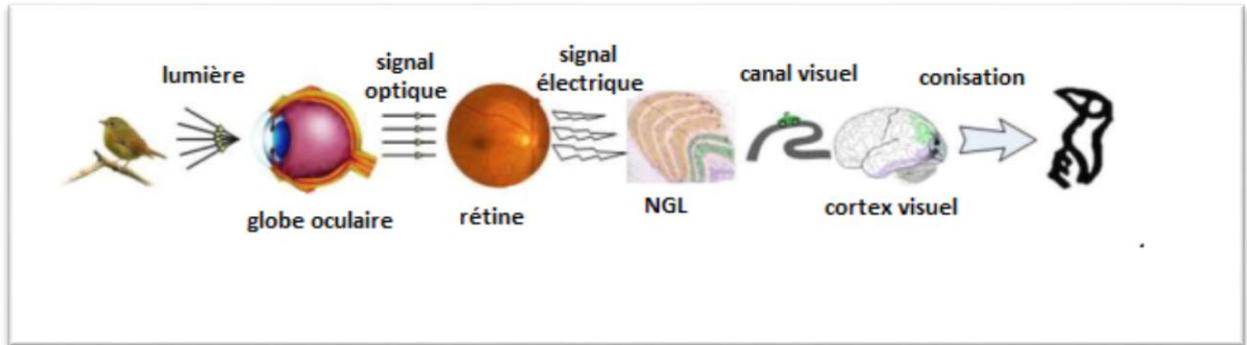


Figure 1 : Traitement de l'information visuelle dans le système visuel humain [3].

Les études de physiologie visuelle montrent que le traitement de l'information visuelle comporte quatre aspects le traitement optique, le traitement de la rétine, le traitement géniculé latéral et traitement du cortex visuel [3]. Le traitement des informations visuelles dans le système HVS est illustré à la Figure 1.

#### 4. La Structure des Yeux

Les yeux sont les organes humains permettant de produire une image du monde réel au cerveau. Les primates, y compris les humains, ont une vision sophistiquée utilisant deux yeux, appelée vision binoculaire. Les signaux visuels passent de chaque œil le long des millions de fibres du nerf optique vers le chiasma optique, où certaines fibres nerveuses se croisent desorte que les deux côtés du cerveau reçoivent des signaux des deux yeux [2]. La moitié gauche des deux rétines se projette sur le cortex visuel gauche et la moitié droite sur le cortex visuel droit.

##### 4.1. Composants de l'œil

Dans les deux globes oculaires se constitue trois enveloppes emboîtées la sclérotique, la choroïde et la rétine.

##### 4.1.1. La sclérotique

La sclérotique est une membrane blanche et opaque, très résistante, de structure tendineuse et d'épaisseur de 1 à 2 mm, qui forme le « blanc » de l'œil. Elle représente la partie blanche et opaque de l'œil.

##### 4.1.2. La choroïde

La choroïde est située à l'intérieur de la sclérotique ; c'est une membrane richement vascularisée et imprégnée d'un pigment noir, la mélanine. Vers l'avant, la choroïde donne la naissance à l'iris, un disque coloré percé en son centre par un orifice, la pupille, dont le diamètre varie en fonction de l'intensité lumineuse qui pénètre dans l'œil.

### 4.1.3. La rétine

La rétine est située à la surface interne de l'œil. Elle contient des photorécepteurs qui ne perçoivent qu'une partie du spectre de la lumière visible. Il y a environ 125 millions de photorécepteurs dans chaque œil et ils sont spécialisés dans la transformation de la lumière en signaux électriques. Ces photorécepteurs sont classés en bâtonnets et cônes.

Les bâtonnets sont plus sensibles à la lumière tamisée et ne véhiculent pas de couleur, elles permettent de voir dans les faibles conditions d'éclairage (vision nocturne), tandis que les cônes fonctionnent à la lumière vive et sont responsables des détails aigus, de la vision en noir et blanc et de la vision des couleurs, elles sont donc adaptées à l'éclairage de la lumière du jour (vision diurne).

Les photorécepteurs sont connectés à des cellules horizontales et bipolaires [4]. Une cellule bipolaire peut recevoir des connexions directement des photorécepteurs ou indirectement via la cellule horizontale. La sortie des cellules bipolaires est intégrée dans les cellules ganglionnaires. Ces cellules entières sont responsables de la conversion photoélectrique et de la transmission des informations. La structure de l'œil est illustrée à la figure 2.

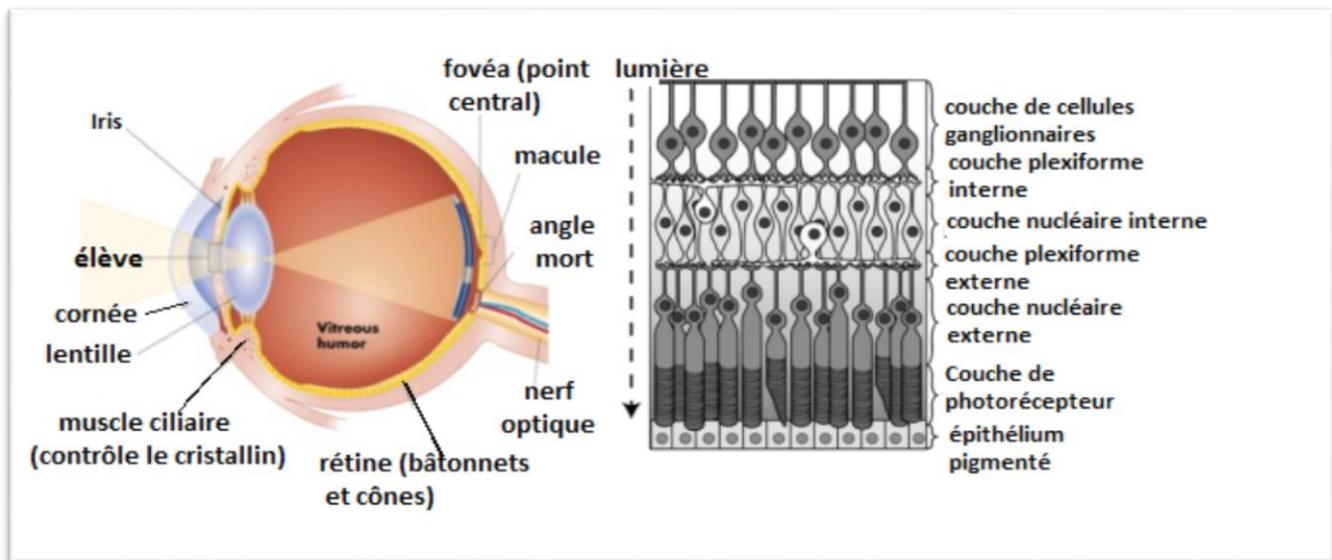


Figure 2 : (a) La structure de l'œil ; (b) La structure laminaire de la rétine qui montre les cinq premières couches. La lumière doit traverser ces couches pour atteindre les photorécepteurs situés à l'arrière de la rétine [5].

## 5. La perception visuelle

La perception visuelle regroupe les mécanismes mis en œuvre pour la réception et la cognition de stimuli visuels. La partie réceptive se charge de capter et d'organiser les informations visuelles en provenance de l'environnement, alors que la partie cognitive se charge de l'interprétation de ces informations. [6]

## 6. La perception des couleurs

### 6.1. Couleur

C'est l'impression produite sur l'œil par les diverses radiations constitutives de la lumière. La couleur est la perception que nous avons des différentes longueurs d'onde qui constituent la lumière visible. Cet ensemble de longueurs d'onde qu'on appelle le spectre de la lumière s'étend du violet (longueur d'onde = 400 nanomètres) au rouge (longueur d'onde = 700 nanomètres). Au-delà de ces longueurs d'onde, la lumière devient invisible et on entre dans le domaine de l'ultraviolet (rayons responsables du bronzage) et dans l'infrarouge ou rayonnement calorifique.

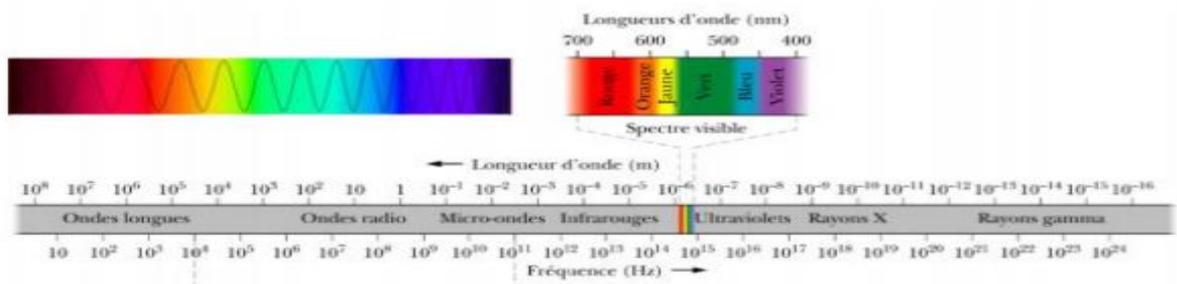


Figure 3 : Spectre des ondes électromagnétiques et différents domaines d'application (extrait du cours de Physique Générale PhyG II-7).

Si on parle de couleur, on parle de la lumière, dont la présence est nécessaire pour la discrimination des couleurs. La lumière est décrite par une onde électromagnétique de haute fréquence et par conséquent de courte longueur d'onde. La lumière visible correspond aux radiations électromagnétiques dont la longueur d'onde dans le vide est comprise entre 400 nm et 800 nm.

## 7. Les systèmes de couleurs

L'idée que le système de vision humaine est trichromique remonte presque simultanément à James Clerk Maxwell (1855) et à Herman Gunther Grassman (1853). Ces lois, connus sous le nom de principe de composition linéaire ou de lois de Grassman ont posé les principes de la vision humaine des couleurs. Elles sont au nombre de trois

- **Trichromie** : Toute sensation colorée peut être reproduite par un mélange additif de trois couleurs primaires convenablement choisies.

- **Additivité** : La sensation colorée provoquée par un mélange additif de deux ou plusieurs lumières colorées est égale à la somme des intensités des primaires correspondant à chacune des lumières.
- **Dilatation** : Si une couleur baisse ou augmente en intensité, il faut et il suffit pour retrouver la même sensation visuelle d'augmenter dans la même proportion les composants des trois primaires qui la composent.

Un système se définit par le choix des primaires utilisés et du blanc de référence qui fixe leurs valeurs unitaires.

Il existe 02 manières de classer les couleurs

- Selon une approche physique RVB, CMJ
- Selon une approche perceptuelle TSL, CIELAB, CIELUV. [x11]

Avant de définir ces systèmes on va d'abords définir quelques termes

- **Teinte** (en anglais hue) c'est le nom qu'on utilisera pour designer la couleur « vert », « mauve », « orange » ...etc. Il est associé à une longueur d'onde.



- **Saturation** « intensité » de la couleur

Varie entre 0 et 100 %, est parfois appelé « pureté ». Plus la saturation d'une couleur est faible, plus l'image sera « grisée » et plus elle apparaîtra fade, il est courant de définir la « désaturation » comme l'inverse de la saturation.



- **Luminance**

C'est la mesure de l'intensité lumineuse de la couleur, qui doit varier entre le noir absolu et le blanc.



### 7.1. Le système colorimétrique RVB (ou RGB en anglais)

C'est un système de codage informatique des couleurs le plus proche du matériel.

On constate 2 synthèses de couleurs

- **Synthèse additive** c'est pour décrire les couleurs d'un écran d'ordinateur ; les couleurs perçus de transmission de lumière sont Rouge + Vert + Bleu = Blanc.

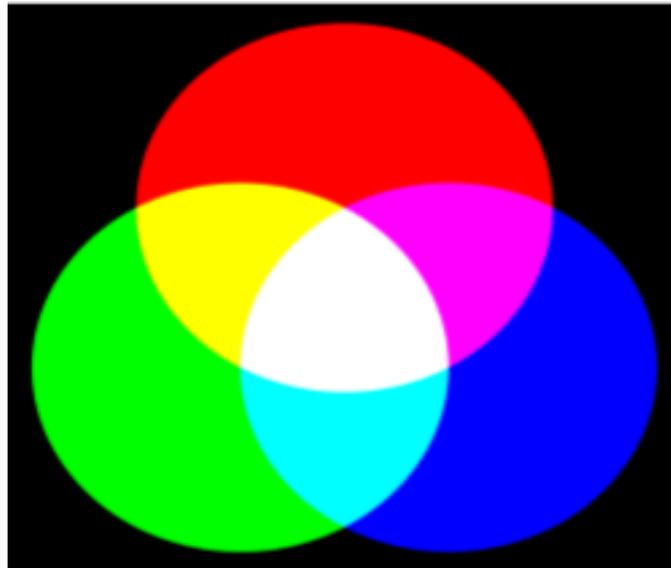


Figure 4 : Couleurs Additives (Les couleurs primaires rvb)

- **Synthèse soustractive** c'est pour décrire les couleurs des matériaux d'impression comme l'imprimante ; les couleurs perçus de réflexion de lumière sont Cyan + Magenta + Jaune = Noir.

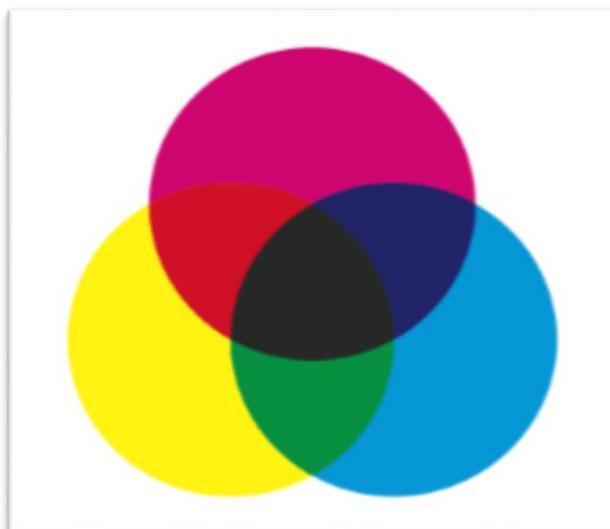


Figure 5 : couleurs soustractives (couleurs primaire CMJ « film, imprimante »)

Chaque stimulus de couleur « C » est ainsi représenté par un point qui définit un vecteur couleur. Les coordonnées de ce vecteur sont les composantes trichromatiques  $R_c$ ,  $V_c$  et  $B_c$ .

Les points correspondant à des stimuli de couleur dont les composantes trichromatiques sont positives sont contenus dans un cube, connu sous le nom de cube des couleurs.

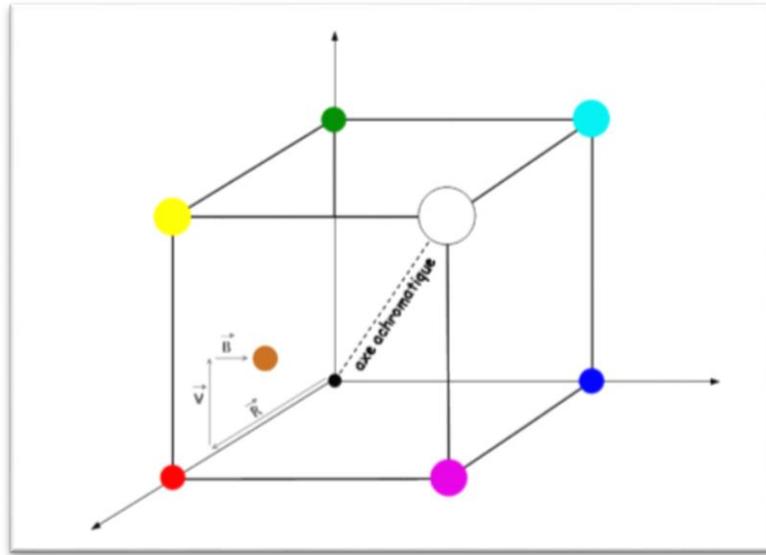


Figure 6 : Système RVB.

L'origine correspond au noir tandis que le blanc de référence est défini par le mélange unitaire des trois primaires. La droite passant par les points Noir (0,0,0) et Blanc (1,1,1) est appelée axe des gris, axe des couleurs neutres ou encore axe achromatique.

Les points de cette droite représentent des nuances de gris allant du noir au blanc. Elle a pour équation  $R_c = G_c = B_c$ .

## 7.2. Le système CIE XYZ

L'espace XYZ (CIE 1931) est défini à partir d'une transformation linéaire sur l'espace RGB, tel que toutes les couleurs du spectre visible soient contenues dans le triangle xyz.

L'espace CIE XYZ introduit la notion de luminance, intensité lumineuse subjective indépendante de la couleur, donnée par la composante Y.

Il utilise deux autres composantes X et Z, choisies de telle manière qu'elles prennent toujours des valeurs positives pour décrire les couleurs visibles.

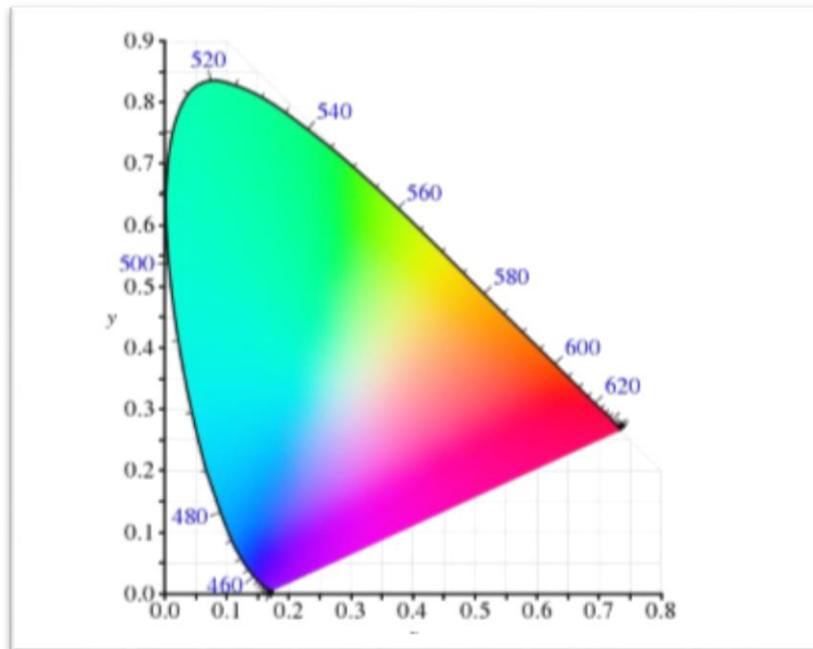


Figure 7 : Diagramme de chromaticité CIE ( $x, y$ ). La coordonnée trichromatique  $x$  est en abscisse et  $y$  en ordonnée.

Le triplet  $x, y, z$  correspond une seule couleur perçue par l'œil de l'observateur standard défini par la formule

$$S = x \cdot X + y \cdot Y + z \cdot Z,$$

Avec  $x, y, z$  représentant les coordonnées trichromatiques et  $X, Y, Z$  représentant les primaires

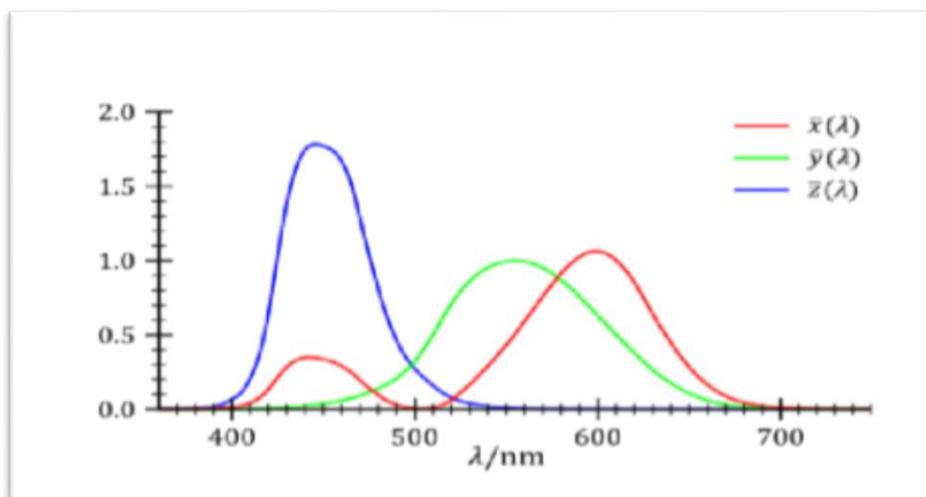


Figure 8 : CIE XYZ (fonction d'observation d'un standard).

### 7.3. Le système TSL /TSV (Teinte Saturation Luminance / Teinte Saturation Valeur)

Ce système est représenté par la Teinte, Saturation, et la Luminance.

D'autre représentation par Teinte, Saturation, et valeur.

- La Valeur est la « brillance » de la couleur ; elle varie entre 0 et 100 %.

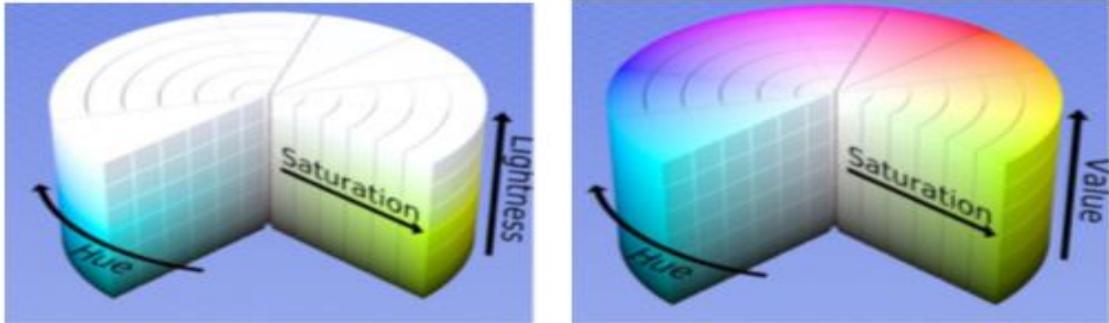


Figure 9 : TSL /TSV (Teinte Saturation Luminance / Teinte Saturation Valeur)

### 7.4. L'espace colorimétrique absolu CIE L\*a\*b 1976

C'est un modèle de représentation des couleurs développé en 1976 par le CIE ; elle contient environ 4.2 millions de couleurs

L'espace L\*a\*b est une tentative

- Pour linéariser la réponse logarithmique de l'œil.
- Préciser sa tolérance tout en tenant compte de sa grande sensibilité à la luminance.
- Et en évitant de représenter les couleurs dans un espace déformé comme c'est le cas dans tous les autres modèles dérivés de l'espace CIE.

Dans l'acronyme L\*a\*b, le signe \* représente la dérivée des valeurs L, a et b, et en pratique

- L\* représente la luminance.
- a\* est un point sur l'axe vert-rouge.
- b\* un point de l'axe jaune-bleu, leur produit définissant un point dans l'espace colorimétrique.

Le système CIELAB 1976 possède deux modes de représentation

- Représentation en coordonnées rectangulaires L\*a\*b\*.
- Représentation en coordonnées cylindriques L\*C\*h°.

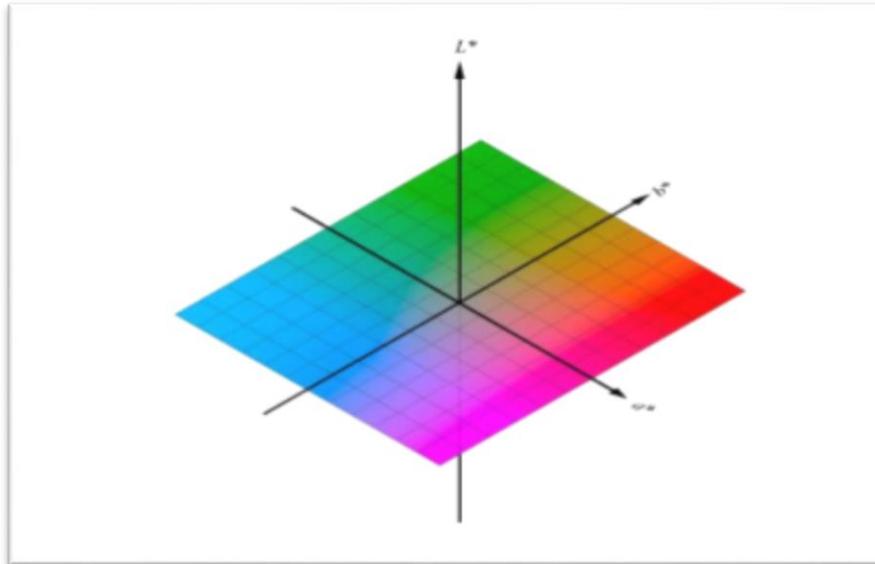


Figure 10 : Espace  $l^* a^* b^*$ .

## 8. Conclusion

Dans ce chapitre nous avons essayé, d'introduire les éléments de base nécessaires pour comprendre la problématique à laquelle nous allons être confrontés par la suite. Nous avons ainsi vu que la perception visuelle ne se base pas seulement sur l'organe de l'œil mais également sur le cerveau.

En effet, la vision n'est pas seulement liée aux yeux. Les informations de ces derniers sont transmises au cerveau qui les traite dans différentes aires visuelles reliées par des voies spécifiques.

C'est pour cela, que nous avons défini les éléments principaux de l'œil, et que nous avons introduit le principe de la perception et en particulier la manière de percevoir la couleur et comment la couleur est modélisée.

# **Chapitre 02 : Classification des modèles de l'attention visuelle**

## 1. Introduction

L'attention peut être abordée selon différentes perspectives. Celles-ci reposent sur des classifications de l'attention, qui se fait en fonction de l'approche choisie, ainsi les classifications utilisées recouvrent des notions plus ou moins différentes et permettent de répondre à des questions variées et renvoient la représentation de l'attention visuelle à des façons différentes.

En effet, il semble qu'en fonction des problématiques des chercheurs, les questions ont évolués. Nous nous attarderons ici sur trois questions : les processus de l'attention peuvent se différencier en terme de comportement moteur (vont-ils être associés à une action ?), en terme de traitement (l'information sera-t-elle traitée différemment en fonction de ceux-ci ?), en terme de sélection de l'information (quelle information vont-ils sélectionner ?).

Nous allons donc présenter, de manière non exhaustive, les distinctions principales de l'attention visuelle, qui vont nous permettre de situer notre travail et de justifier le choix taxonomique fait par la suite

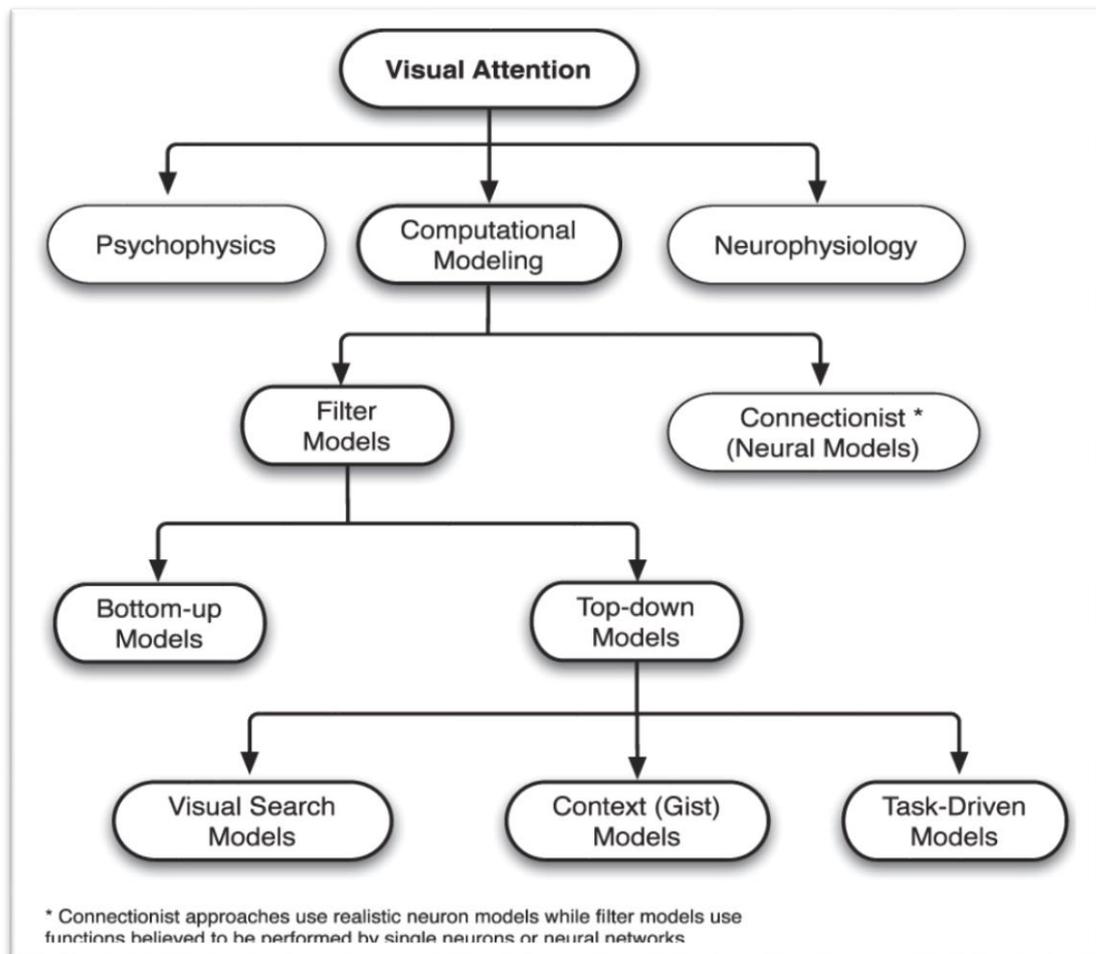


Figure 11 : Classification des modèles de l'attention visuelle

## 2. Attention visuelle

L'attention ou la discrimination visuelle caractérise intuitivement les parties d'une scène qui se détachent par rapport aux parties voisines et attirent notre attention. Examiner les mécanismes du système visuel humain dans la sélection des régions d'intérêt est un problème de recherche important et fondamental en neurosciences et en psychologie.

Pour le système visuel humain, il est impossible de percevoir simultanément la scène entière dans le même acte sensoriel. Seule une petite région de la scène est analysée en détail à chaque instant. La région actuellement visitée ne sera pas toujours la même que celle fixée par les yeux et le regard passera à la prochaine région intéressante. Ce processus se poursuit jusqu'à ce que la scène entière soit perçue. L'ordre dans lequel la scène est analysée est déterminé par le mécanisme de l'attention sélective [7].

Le schéma précédent (figure 11), montre une classification des différents modèles de l'attention visuelle. Nous allons expliquer dans ce qui suit ces modèles :

### 2.1. Psychophysique

La psychophysique est une branche de la psychologie expérimentale qui cherche à déterminer les relations quantitatives qui existent entre un stimulus physique et la perception qu'on en a.

La psychophysique s'intéresse aux sens physiologiques tels que la vue, l'ouïe, le toucher (plus rarement l'odorat ou le goût) mais aussi à des sensations comme la perception du temps ou du mouvement.

### 2.2. Neurophysiologie

La neurophysiologie est l'étude des fonctions du système nerveux, reposant sur tous les niveaux de description, du niveau moléculaire jusqu'au niveau le plus intégré des réseaux de neurones. C'est une science pluridisciplinaire, au carrefour d'autres sciences biologiques (neuroanatomie neuroendocrinologie, neuropharmacologie....) et comportementales (neuropsychologie, psychiatrie...).

### 2.3. Modélisation computationnelle

Les modèles computationnels d'attention, dans la mesure où ils prédisent relativement bien l'attractivité visuelle de différents emplacements d'une œuvre d'œuvre, peuvent offrir une solution plus simple et plus rentable au problème de l'élaboration de critères d'évaluation rapides, objectifs et impartiaux. L'hypothèse est que les lieux marqués comme très saillants par le modèle devraient attirer le regard d'une majorité de clients potentiels. Une meilleure conception est donc celle qui transmet les informations qui importent à l'annonceur à ces emplacements [8].

### 2.3.1. Modèles neuronaux

Les modèles informatiques de la cognition axés sur l'utilisation des réseaux de neurones. Ces architectures ont été inspirées par des recherches sur le fonctionnement du calcul dans le cerveau et les travaux ultérieurs ont abouti à des modèles cognitifs à la saveur distincte. Le traitement est caractérisé par des modèles d'activation sur des unités de traitement uniques connectées ensemble dans des réseaux complexes. La connaissance est stockée dans la force des liens entre les unités. C'est pour cette raison que cette approche de la compréhension de la cognition a pris le nom de connexionnisme [9].

### 2.3.2. Modèles de filtres

Une distinction importante entre les modèles est de savoir s'ils s'appuient sur des influences ascendantes (Bottom-up), des influences descendantes (Top-down) ou une combinaison des deux. Les indices de Bottom-up sont principalement basés sur les caractéristiques d'une scène visuelle (déterminée par un stimulus), alors que les indices de Top-down (déterminés par un objectif) sont déterminés par des phénomènes cognitifs tels que les connaissances, les attentes, les récompenses et les objectifs actuels.

#### 2.3.2.1. Modèles de Top-down

L'attention descendante (Top-down) est lente, axée sur les tâches, volontaire et en boucle fermée. L'un des exemples les plus célèbres de guidage par le haut de l'attention est celui de Yarbus en 1967, qui a montré que les mouvements oculaires dépendaient de la tâche actuelle associée à l'expérience suivante il était demandé aux sujets de regarder la même scène (une pièce avec une famille). Et un visiteur inattendu entrant dans la salle) dans différentes conditions (questions) telles que «estimer les circonstances matérielles de la famille», «quel âge ont les gens?» ou simplement examiner librement la scène.

Les mouvements oculaires différaient considérablement dans chacun de ces cas. Les modèles ont exploré trois sources principales d'influences top-down en réponse à cette question comment décider où regarder ? Certains modèles traitent de la recherche visuelle dans laquelle l'attention est attirée sur les caractéristiques d'un objet cible que nous recherchons. D'autres modèles étudient le rôle du contexte de la scène ou de l'essentiel pour contraindre les emplacements que nous examinons. Dans certains cas, il est difficile de dire précisément où et ce que nous examinons car une tâche complexe régit la fixation des yeux, par exemple, la conduite. Tandis que, en principe, les demandes de tâches sur l'attention englobent les deux autres facteurs, dans la pratique, les modèles se sont concentrés sur chacun d'eux séparément. La disposition des scènes a également été proposée comme source d'attention descendante et est considérée ici avec le contexte de la scène.

##### 2.3.2.1.1. Modèles de recherche visuelle (visual search models)

Les modèles attentionnels ont généralement été validés contre les mouvements oculaires d'observateurs humains. Les mouvements oculaires transmettent des informations importantes sur les processus cognitifs tels que la lecture, la recherche visuelle et la perception de la scène. En tant que tels, ils sont souvent traités comme des substituts des changements d'attention. Par exemple, dans la perception des scènes et la recherche visuelle, lorsque le stimulus est plus encombré, les fixations deviennent plus longues et les saccades

plus courtes. La difficulté de la tâche (lire, par exemple, pour comprendre, plutôt que pour lire l'essence, ou rechercher une personne dans une scène, ou regarder la scène pour un test de mémoire) influence évidemment le comportement des mouvements oculaires. Bien que les modèles de prévision de l'attention et des mouvements oculaires soient souvent validés par rapport à des données oculaires, il existe de légères différences dans la portée, les approches, les stimuli et le niveau de détail. Les modèles de prédiction des mouvements oculaires (programmation de saccades) tentent de comprendre les fondements mathématiques et théoriques de l'attention. Certains exemples incluent les processus de recherche (par exemple, la théorie de la recherche optimale), les modèles de maximisation de l'information [10].

Il faut noter que les mouvements oculaires ne racontent pas toujours toute l'histoire et qu'il existe d'autres métriques pouvant être utilisées pour l'évaluation du modèle. Par exemple, la précision dans le signalement correct d'un changement dans une image (c.-à-d. Aveuglement par la recherche) ou dans la prédiction des éléments qui retiennent l'attention dont vous vous souviendrez montre des aspects importants de l'attention qui ne sont pas pris en compte par la seule analyse des mouvements oculaires.

De nombreux modèles d'attention en recherche visuelle ont également été testés en estimant avec précision les temps de réaction (RT) (par exemple, RT / définition des pentes dans les tâches de recherche contextuelle et de recherche de conjonction) [10].

#### 2.3.2.1.2. Modèles de Contexte (context (Gist) models)

Après une brève présentation d'une image (80 ms ou moins), un observateur est capable de rendre compte des caractéristiques essentielles d'une scène. Cette représentation très grossière d'une scène, dite "élément essentiel", ne contient pas beaucoup de détails sur des objets individuels, mais peut fournir suffisamment d'informations pour discriminer les scènes grossières (par exemple, en intérieur ou en extérieur).

Il est important de noter que l'essentiel ne révèle pas nécessairement la catégorie sémantique d'une scène. Chun et Jiang ont montré que les cibles apparaissant dans des configurations répétées par rapport à certains objets d'arrière-plan (distracteur) étaient détectées plus rapidement. Les associations sémantiques entre les objets d'une scène (par exemple, un ordinateur est souvent placé sur un bureau) ou des indices contextuels jouent également un rôle important dans la direction des mouvements oculaires [10].

Les représentations Gist sont devenues de plus en plus populaires dans la vision par ordinateur car elles fournissent des informations globales et discriminantes, utiles pour de nombreuses applications telles que la recherche dans les jeux de données de scènes à grande échelle disponibles aujourd'hui, en limitant la recherche aux emplacements susceptibles de contenir un objet d'intérêt, l'achèvement de la scène, et la modélisation de l'attention du Top-down). On peut donc constater que les recherches dans ce domaine sont potentiellement très prometteuses [10].

### 2.3.2.1.3. Modèles axés sur les tâches (Task-driven model)

La tâche a une forte influence sur le déploiement de l'attention. Nous pouvons confirmer que les scènes visuelles sont interprétées selon les besoins pour répondre aux demandes de tâches. Ainsi Hayhoe et Ballard ont montré qu'il existait une relation étroite entre la cognition visuelle et les mouvements oculaires lorsqu'il s'agissait de tâches complexes.

Il a été constaté que les sujets effectuant une tâche guidée visuelle dirigeaient la majorité des ajustements vers des emplacements pertinents pour la tâche. Il est souvent possible de déduire l'algorithme qu'un sujet a en tête à partir du schéma de ses mouvements oculaires. Par exemple, dans une tâche de "copie de bloc" dans laquelle les sujets devaient répliquer un assemblage de blocs de construction élémentaires, l'algorithme des observateurs pour effectuer la tâche était révélé par des schémas de mouvements oculaires.

Les sujets ont d'abord sélectionné un bloc cible dans le modèle pour vérifier la position du bloc, puis ont fixé l'espace de travail pour placer le nouveau bloc à l'emplacement correspondant. D'autres recherches ont étudié des comptes rendus de haut niveau sur le comportement du regard dans des environnements naturels pour des tâches telles que la fabrication de sandwich, la conduite, le cricket et la marche Sodhi et al ont étudié l'impact des distractions au volant, telles que le réglage de la radio ou la réponse au téléphone, sur les mouvements des yeux [10].

### 2.3.2.2. Modèles de Bottom-up

Les régions d'intérêt qui attirent notre attention de manière ascendante (Bottom-up) doivent être suffisamment distinctives par rapport aux entités environnantes. Ce mécanisme attentionnel est également appelé exogène, automatique, réflexif ou périphérique.

L'attention ascendante (Bottom-up) est rapide, involontaire et très probablement rétroactive, Un exemple type d'attention ascendante consiste à regarder une scène avec une seule barre horizontale parmi plusieurs barres verticales est immédiatement attiré par la barre horizontale.

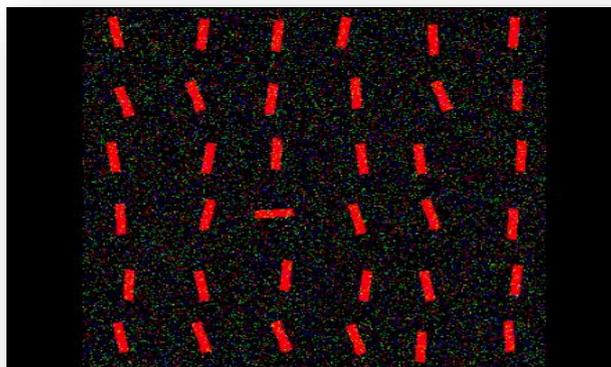


Figure 12 : Approche ascendante (Bottom-up)

Elle est liée à la notion de saillance qui définit une région saillante d'une image comme celle qui se distingue de son voisinage par certains attributs visuels, Cette région saillante attire l'attention. La saillance s'appuie donc uniquement sur le contenu de l'image et les régions saillantes d'une image sont donc les mêmes quel que soit l'observateur.

#### 2.3.2.2.1. Carte de Saillance

La carte de saillance est une carte topographique arrangée qui représente la saillance visuelle d'une scène visuelle correspondante.

Les différents éléments visuels qui contribuent à la sélection attentive d'un stimulus (couleur, orientation, mouvement, ...) sont combinés en une seule carte orientée topographique, la carte de saillance intègre l'information normalisée à partir des cartes de caractéristiques

Individuelles vers une mesure globale de visibilité.

La carte de saillance comprend les éléments suivants

- Une représentation anticipée composée d'un ensemble des cartes de caractéristiques calculées en parallèle.
- Une carte de saillance topographique où chaque emplacement encode la combinaison des propriétés sur toutes les cartes de caractéristiques.
- Une cartographie sélective des propriétés d'un seul emplacement visuel représentée d'une manière non-topographique centrale, à partir de la carte de saillance.
- Un réseau WTA (Winner-take-all) implémente le processus de sélection.
- Inhibition de l'emplacement choisi, qui provoque un décalage automatique vers l'emplacement prochain le plus visible.

Les cartes de caractéristiques encodent la visibilité avec une dimension particulière des caractéristiques, et la carte de saillance combine les informations de chaque carte de caractéristiques. La saillance à un emplacement donné est déterminée par le degré de différence entre cet emplacement et ces voisins [11].

La figure 13 illustre les différents éléments de la carte de saillance par le premier modèle d'attention visuelle de Koch & Ulfman.

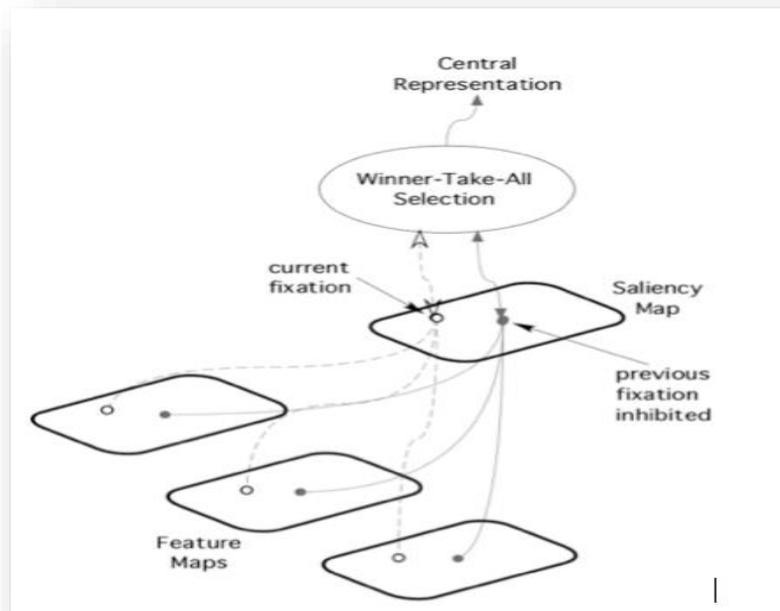


Figure 13 : Modèle de carte de saillance Koch & Ullman. (Extrait de [Koch et Ullman, 1987])

### 3. Modèles d'attention

Dans cette partie, les modèles d'attention sont expliqués en fonction de leur mécanisme d'obtenir la saillance. Certains modèles se répartissent en plusieurs catégories. Ces modèles ont été implémentés et peuvent traiter des images numériques arbitraires et donner les cartes de saillance correspondantes.

#### 3.1. Modèles Cognitifs

La plupart des modèles d'attention sont directement ou indirectement inspirés par le concept cognitif. Le modèle d'*itti* est le modèle de base. Le modèle *Itti & al* [12] utilise trois caractéristiques la couleur, l'intensité et l'orientation. Ce modèle est la base des modèles les plus récents et la référence standard pour la comparaison [11].

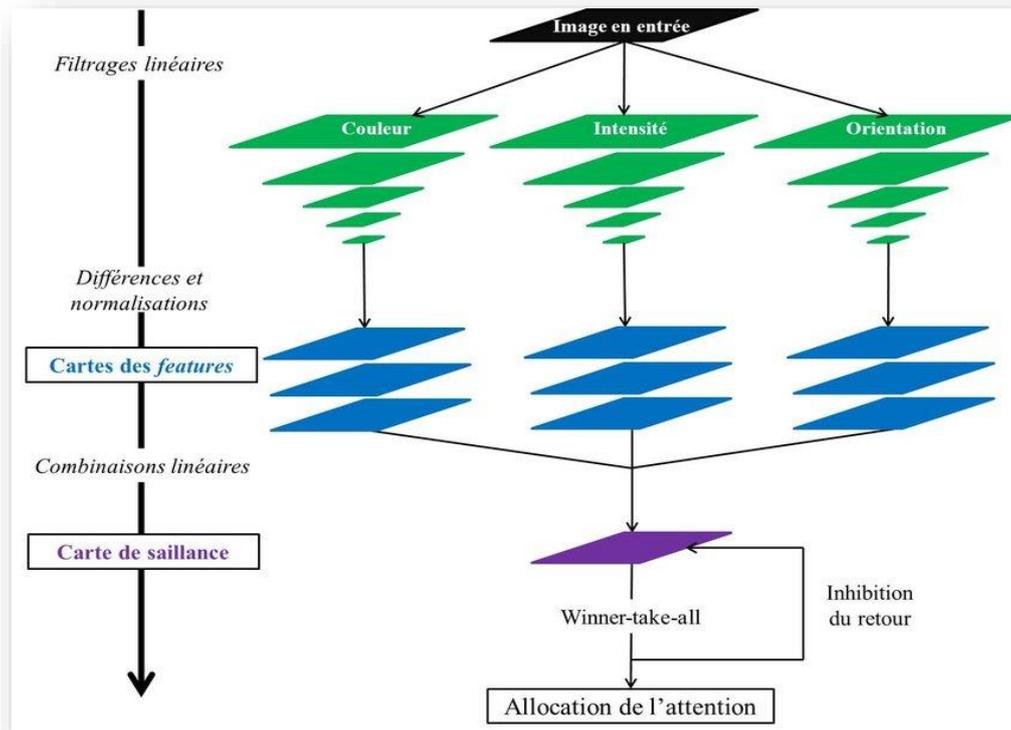


Figure 14 : Le modèle décrit par Itti et al. (1998) puis par Itti & Koch (2001).

Le modèle décrit par Itti et al. (1998) puis par Itti & Koch (2001). L'image donnée en entrée est décomposée en fonction des trois principales features (l'orientation, l'intensité et la couleur) et en fonction des zones de l'espace, à l'image des champs réceptifs des neurones. Le modèle produit ensuite des cartes de features sur la base des informations précédemment recueillies, puis, par combinaison de ces cartes, à une carte de saillance unique. En suivant le principe de winner-take-all (i.e. le point le plus saillant est choisi), le modèle prédit la prochaine position de l'attention visuelle. Puis, an d'éviter de xer ce même point indéniement, l'inhibition du retour permet d'inhiber ce point an que la position de l'attention change.

### 3.2. Modèles Bayésiens

La modélisation bayésienne est utilisée pour combiner les preuves sensorielles et les contraintes préalables. Dans ces modèles, les connaissances préalables (contexte de la scène, points principales...) et les informations sensorielles (caractéristiques de la cible) sont combinés d'une manière probabiliste selon la règle de Bayes, pour détecter les objets d'intérêt [11].

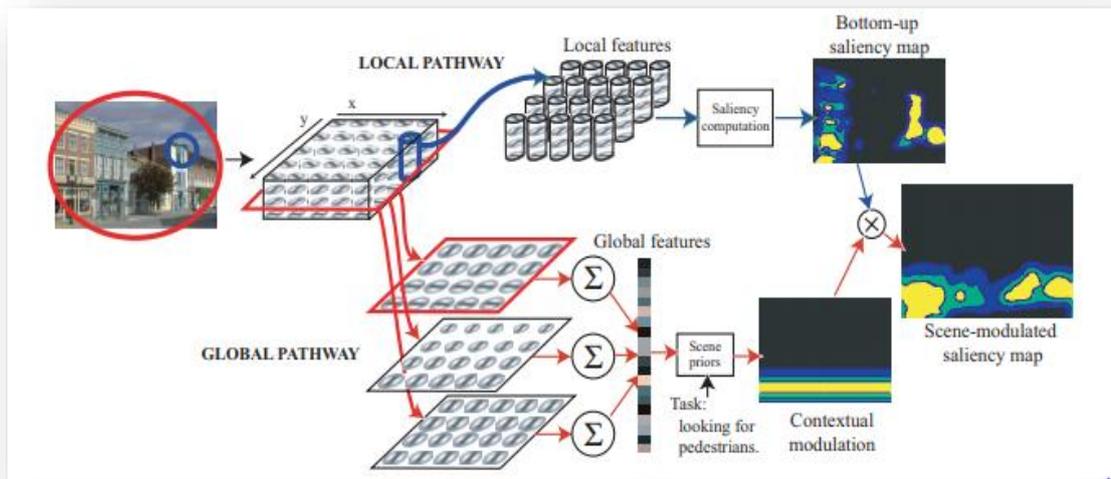


Figure 15: Modèle Bayésien

### 3.3. Modèles de Théorie de Décision

L'interprétation décision-théorique indique que les systèmes de perception évoluent pour produire des décisions sur les états de l'environnement entourant, qui sont optimales dans un sens de décision théorique (probabilité minimale d'erreur). Le point fondamental est que l'attention visuelle devrait être entraînée par optimalité par rapport à la tâche finale [11].

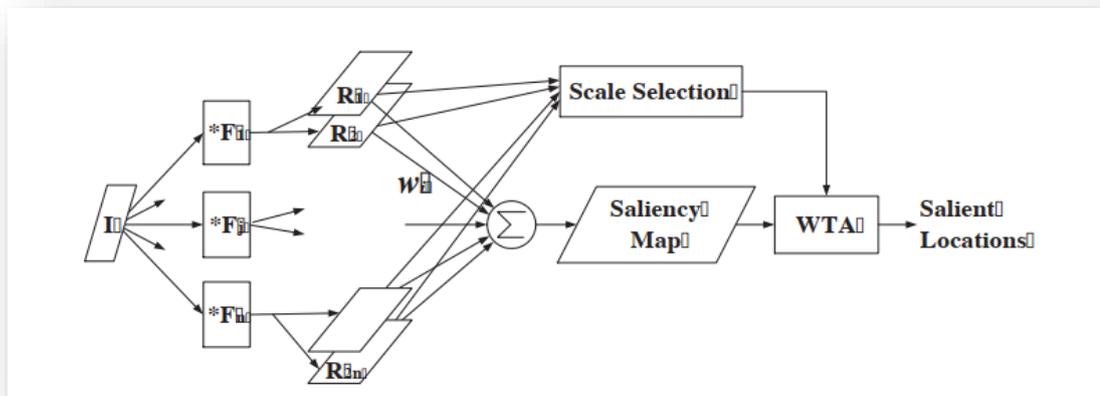


Figure 16 : Modèle de Théorie de Décision

### 3.4. Modèles de Théorie de l'information

Ces modèles sont basés sur l'hypothèse que le calcul de la saillance localisée, sert à maximiser l'information échantillonnée à partir de son environnement. Leur traitement est de sélectionner les parties les plus informatives d'une scène et de négliger le reste [11].

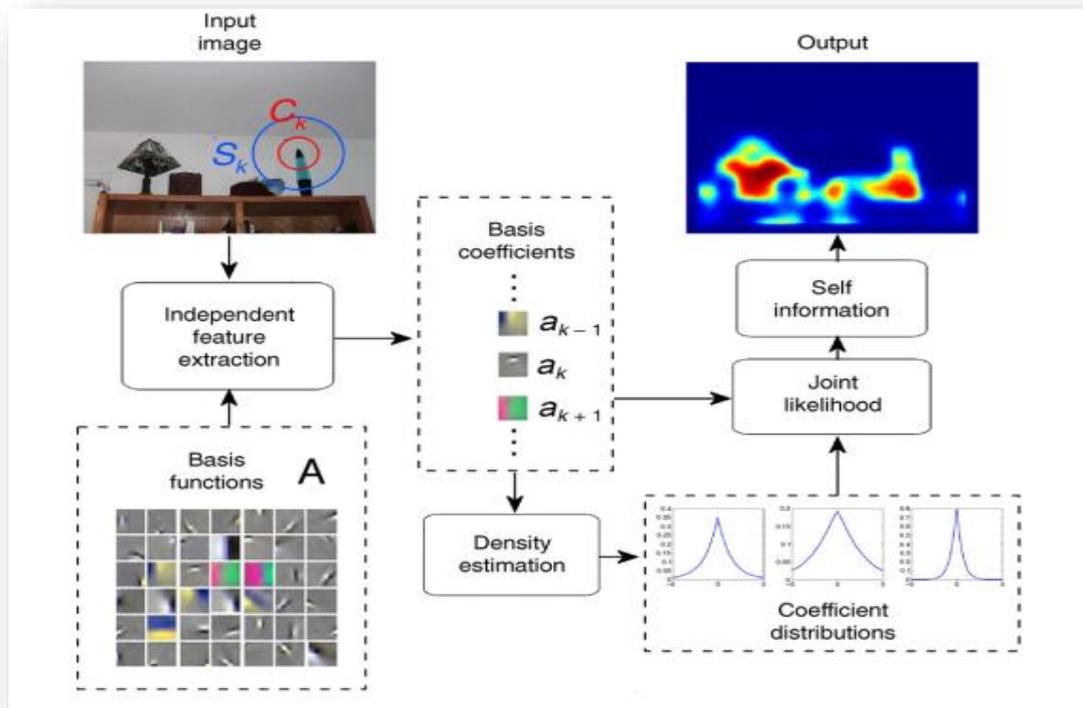


Figure 17 : Modèle de Théorie de l'information

### 3.5. Modèles Graphiques

Un modèle graphique est un modèle probabiliste dont lequel un graphe indique une structure d'indépendance conditionnelle entre des variables aléatoires. Les modèles d'attention dans cette catégorie traitent l'œil comme une série temporelle. Comme ils existent des variables cachées influence la génération du mouvement de l'œil, des approches sont incorporées comme *modèle de markov caché (MMC)*, *Réseau bayésien dynamique (RBN)*, *champs aléatoires conditionnels (CAC)*. Harel & al [Harel et al., 2006] ont introduit le GBVS (Graph-based visual saliency) une saillance visuelle basée sur les graphes. Les cartes de caractéristiques sont extraites à plusieurs échelles spatiales. Une pyramide échelle-espace est dérivée à partir des caractéristiques d'image : intensité, couleur, orientation. Ensuite, un graphe entièrement connecté est construit sur tous les emplacements de la grille de chaque carte de caractéristiques. Les poids entre deux nœuds sont affectés proportionnellement à la similarité des valeurs de caractéristiques et leur distance spatiale. La dissemblance entre deux positions  $(i,j)$  et  $(p,q)$  dans la carte de caractéristiques [11].

Le graphe résultat est traité comme une chaîne de Markov, par la normalisation des poids sortant de chaque nœud à 1, et par définition d'une relation d'équivalence entre les nœuds et les états, et aussi entre les poids et les probabilités de transition. Leur distribution à l'équilibre est adoptée comme des cartes d'activation et de saillance. Les cartes d'activation sont finalement normalisées, et ensuite combinés en une carte globale unique.

### 3.6. Modèles d'analyse Spectrale

Au lieu de traiter une image dans le domaine spatial, les modèles de cette catégorie dérivent la saillance dans le domaine fréquentiel. *Achanta et al* [Achanta et al., 2009] ont implémenté une approche "frequency-tuned" pour détecter les régions saillants en utilisant les caractéristiques de bas niveau (couleur et luminance). Premièrement l'image d'entrée RGB (I) est transformée dans l'espace de couleur CIE Lab. Ensuite la carte de saillance S de l'image I est calculée comme  $S(x; y) = \| I_{\mu} - I_{\omega_{hc}} \|$  où,  $I_{\mu}$  la moyenne arithmétique de vecteur de caractéristiques de l'image,  $I_{\omega_{hc}}$  La version floue de la gaussienne de l'image originale et  $\|.\|$  distance euclidienne norme L2 [11].

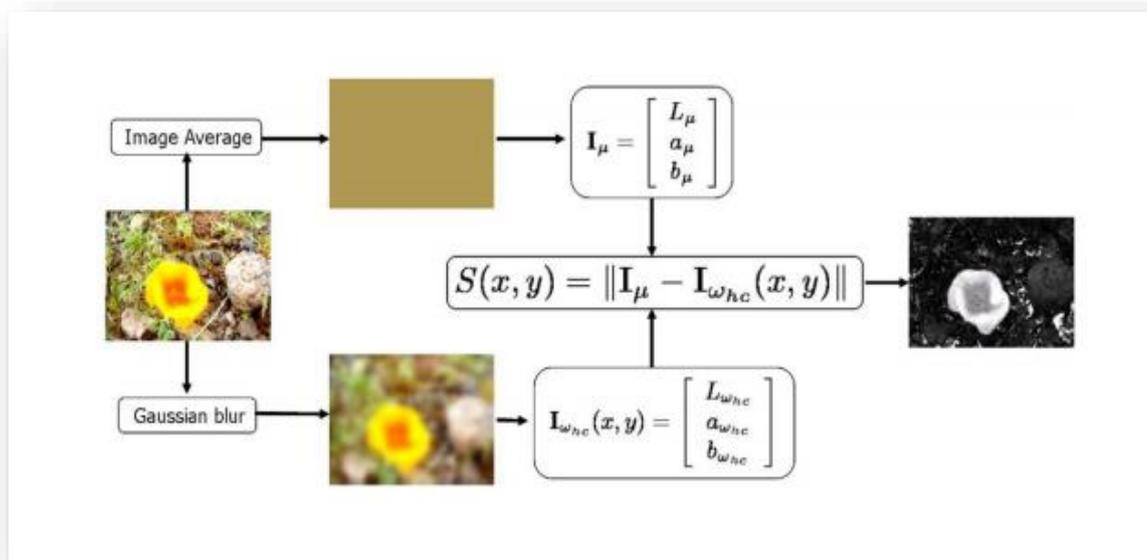


Figure 18 : Modèle d'analyse Spectrale

### 3.7. Modèles de Classification

Les approches d'apprentissage de la machine sont aussi utilisées dans la modélisation de l'attention visuelle par des modèles d'apprentissage à partir des fixations de l'œil enregistrées, ou des régions saillants étiquetées. Noter que ces modèles peuvent qu'ils ne soient pas purement *Bottom-up*, puisque ils utilisent des caractéristiques qui guident l'attention *Top-down* (ex visage, texte, ...).

#### 4. Bilan

Nous avons brièvement présenté les modèles d'attention visuelle en fonction de leurs mécanismes d'obtenir la saillance.

On distingue plusieurs catégories : Les modèles cognitifs, Les modèles bayésiens, Les modèles de théorie de l'information, Les modèles de théorie de la décision, Les modèles graphiques et Les modèles d'analyse spectrale et Les Modèles de classification.

Nous allons dans ce qui suit présenté un bilan qui contient les caractéristiques de chaque modèle ainsi que les avantages et les limites, ce qui nous permettra de guider notre investigation par la suite.

Nous avons résumés ces modèles dans le tableau suivant :

Modèles	Caractéristiques	Avantages	Limites
<b>Modèle Cognitive</b>	Couleur, Intensité, Orientation (Modèle d'Itti).	<ul style="list-style-type: none"> <li>- Elargir notre vision des fondements biologiques de l'attention visuelle.</li> <li>- Aide à comprendre les principes de calcul ou les mécanismes neuronaux de ce processus.</li> <li>-Aide à la reconnaissance des objets.</li> </ul>	- les tests effectués ne sont pas assez satisfaisants, en particulier pour une application telle que la détection de classes d'objet. Lorsqu'aucune des trois caractéristiques d'un objet ne ressort assez, il n'apparaît pas saillant
<b>ModèleBayésien</b>	Combinaison des Connaissances préalable (contexte de la scène, points principales) et des Informations sensorielles (caractéristiques de la cible)	<ul style="list-style-type: none"> <li>- La capacité à apprendre les données.</li> <li>- La capacité à unifier de nombreux facteurs dans un manoir fondé sur des principes.</li> <li>- Les modèles bayésiens peuvent tirer parti des statistiques de scènes naturelles ou d'autres caractéristiques qui attirent l'attention.</li> </ul>	- Calcul des probabilités
<b>Modèle Théorique de Décision</b>	Des filtres de taille fixée obtenus après un entraînement sur la base de données choisi	<ul style="list-style-type: none"> <li>-Les modèles de théorie de la décision ont connu un grand succès dans les applications de vision par ordinateur telles que la classification.</li> <li>- Ils offrent une grande précision dans la prédiction de la fixation.</li> </ul>	- les résultats dépendent de la base d'apprentissage, des taux intéressants sont trouvés lorsque les images tests sont proches de la base. Dans le cas contraire les résultats se dégrade

		- Simple à implémenter	
<b>Modèle Théorie de L'informati on</b>	utilise la mesure <i>auto-informations</i> de Shannon pour calculer la saillance des régions de l'image (La couleur).	- Modèles Théorie de L'information sont simples à expliquer et à mettre en œuvre.	-Problème de dimensionnalité
<b>Modèle Graphique</b>	Intensité, Couleur, Orientation	-Les modèles graphiques pourraient être considérés comme une version généralisée des modèles bayésiens.  -Permet de modéliser des mécanismes d'attention plus complexes sur l'espace et le temps.  - Donne un bon pouvoir de prévision.	-La complexité du modèle, en particulier en ce qui concerne la formation et la lisibilité.
<b>Modèle Spectrale</b>	Couleur, Luminance	-Les modèles d'analyse spectrale sont simples à expliquer et à mettre en œuvre.	-La plausibilité biologique de ces modèles n'est pas très claire.
<b>Modèle de Classificatio n</b>	ils utilisent des caractéristiques qui guident l'attention <i>Top-down</i> (ex visage, texte, ...).	-À mesure que les données disponibles sur les mouvements oculaires augmentent et que les dispositifs de suivi des yeux prenant en charge la collecte de données en masse sont de plus en plus répandus, ces modèles deviennent de plus en plus populaires.	-Dépendants des données, influençant ainsi une comparaison juste des modèles.  - une boîte lente et dans une certaine mesure une boîte noire.

## 5. Conclusion

Nous avons vu dans ce chapitre une classification des modèles de l'attention visuelle, nous avons terminé par une synthèse dressant un bilan sur les caractéristiques, avantages et inconvénients des modèles et nous avons choisi de diriger notre travail vers les modèles à base de la théorie de décision car d'une part offrent une grande précision et d'autre part ils sont simple à implémenter.

# **Chapitre 03 : Technique de l'attention visuelle basée sur la théorie de décision**

## 1. Introduction

Comme nous l'avons indiqué dans le chapitre précédent, nous avons choisi d'implémenter une technique à base de la théorie de décision. Notre choix est motivé par le fait que celle-ci offre des résultats convaincants lorsqu'il s'agit d'images proches de celles existantes dans la base d'apprentissage et que la technique est facile à implémenter.

Notre objectif est donc de mettre en œuvre cette technique et ensuite de comparer les résultats obtenus avec ceux de la technique de base d'Itti et Koch.

Dans ce chapitre nous allons détailler les principales outils utilisés dans la mise œuvre de cette technique.

## 2. Détection de saillance

L'extraction de points saillants à partir d'images est un sujet de recherche depuis au moins quelques décennies. En gros, les détecteurs de saillance peuvent être divisés en trois classes principales :

La première classe, traite le problème comme l'un des problèmes de détection d'attributs visuels. On considère généralement des bords ou des coins, également appelés «points d'intérêt», où leur détection peut être faite à l'aide la structure du mouvement, mais il a également été proposé d'autres attributs visuels de bas niveau tels que les contours. Une limite majeure de ces détecteurs de saillance est qu'ils ne se généralisent pas bien. Par exemple, un détecteur de coin produira toujours une réponse plus forte dans une région à texture forte que dans une région lisse, même si les surfaces texturées ne sont pas nécessairement plus saillantes que les surfaces lisses. Ceci est illustré par l'image de la figure 19 (a) [13].

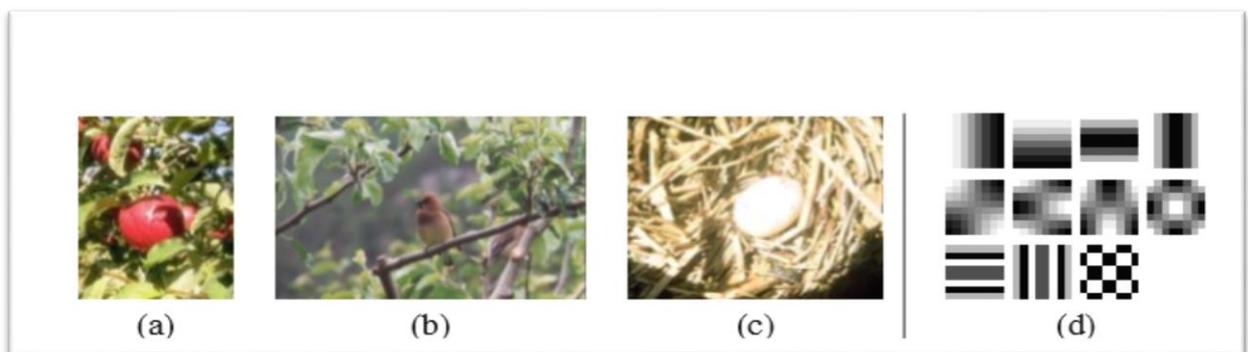


Figure 19 : (a) (b) (c) Divers exemples difficiles pour les détecteurs de saillance actuels. a) pomme suspendue à un arbre, (b) un oiseau dans un arbre, (c) un œuf dans un nid. (d) certaines fonctions de base DCT. De gauche à droite, de haut en bas, les détecteurs de : bords, barres, coins, jonctions en T, points, plages de flux et clés à contrôle

La deuxième classe considère qu'un détecteur de coin répondrait fortement aux régions très texturées des feuilles et des branches d'arbres, mais il n'est pas clair que celles-ci sont plus saillantes que la pomme lisse. Certaines de ces limitations sont abordées dans des publications plus récentes et génériques, formulations de saillance.

La troisième classe définit la saillance comme une complexité d'image. Diverses mesures de complexité ont été proposées dans ce contexte. Lowe [18] mesure la complexité en calculant la variation d'intensité dans une image en utilisant la différence de fonction gaussienne. Sebe [19] mesure la valeur absolue des coefficients d'une décomposition en ondelettes de l'image ; et Kadir [20] s'appuient sur l'entropie de la distribution des intensités locales.

Le principal avantage de ces définitions de la saillance (basées sur les données) est une flexibilité nettement plus grande, comme ils pourraient détecter l'un des attributs de bas niveau ci-dessus (coins, contours, bords lisses, etc.) en fonction de l'image considérée. Cependant, il n'est pas clair que la saillance puisse toujours être assimilée à la complexité. Par exemple, les figures 19 (b) et (c) montrent des images contenant des régions complexes, composées de feuilles en grappes et de paille, qui ne sont pas très saillantes. Au contraire, les régions d'image beaucoup moins complexes contenant l'oiseau ou l'œuf semblent être nettement plus saillantes [13].

### 3. La saillance discriminante

L'intuition de base pour la saillance discriminante est en quelque sorte un «énoncé de l'évidence» : les attributs saillants d'un concept visuel donné sont ceux qui le distinguent le plus de tous les autres concepts visuels potentiellement intéressants. Bien qu'elle soit presque évidente, cette définition s'éloigne de toutes les définitions existantes dans la littérature sur la vision [13].

Premièrement, il fait référence à un «ensemble de concepts visuels d'intérêt possible». Bien que cet ensemble puisse ne pas être bien défini pour tous les problèmes de vision, c'est une composante intrinsèque du problème de reconnaissance : l'ensemble des classes visuelles à reconnaître. Il rend donc la saillance dépendante de l'existence d'un ensemble de classes et, par conséquent, impossible à calculer à partir d'une image isolée. Cela signifie également que, pour un objet donné, différents attributs visuels seront essentiels dans différents contextes de reconnaissance. Par exemple, alors que les contours et la forme seront les plus saillants pour distinguer une pomme rouge d'une voiture rouge, la couleur et la texture seront plus saillants lorsque la même pomme sera comparée à une orange. Toutes ces propriétés semblent être une bonne idée pour la reconnaissance.

Deuxièmement, il a pour objectif principal de distinguer entre les classes. Cela implique que le critère d'optimalité pour la conception des éléments saillants est la discrimination, et donc très différent des critères traditionnels tels que la répétitivité sous des transformations d'image. La robustesse vis-à-vis de ces critères (qui, encore une fois, sont bien justifiés pour le suivi mais ne traitent pas de l'essence du problème de reconnaissance) peut être appris si nécessaire pour atteindre des objectifs discriminants [13].

En raison de cette équivalence entre la saillance et la discrimination, le principe de la saillance discriminante peut facilement être traduit en critère d'optimalité pour la conception d'algorithmes de saillance. En particulier, il est naturellement formulé comme un problème de sélection de caractéristiques optimal: les caractéristiques optimales pour la saillance sont les caractéristiques les plus discriminantes pour le problème de classification (un contre tous) opposant la classe d'intérêt à toutes les classes restantes. Autrement dit, les caractéristiques les plus saillantes sont celles qui séparent le mieux la classe d'intérêts de toutes les autres. Etant donné l'équivalence bien connue entre les caractéristiques et les filtres d'image, ceci peut également être considéré comme un problème de conception de filtres optimaux pour la discrimination [13].

### 3.1. Sélection des caractéristiques évolutives

Dans le contexte des systèmes de reconnaissance évolutifs, la mise en œuvre de la saillance discriminante nécessite 1) la conception d'un grand nombre de classificateurs (autant que le nombre total de classes à reconnaître) au moment de la configuration, 2) le réglage du classificateur chaque fois que de nouvelles classes sont ajoutées ou supprimées au problème. Il est donc important d'adopter des techniques de sélection des caractéristiques qui sont efficace en calcul, réutiliser de préférence le calcul de la conception d'un classificateur à le suivant. La conception de telles méthodes de sélection des caractéristiques est un problème non trivial, que nous avons activement recherché dans le contexte de la recherche sur la sélection des caractéristiques elle-même. Cette recherche a montré que les méthodes théoriques de l'information, basées sur la maximisation de l'information mutuelle entre caractéristiques et étiquettes de classe, ont la propriété intéressante de permettre un contrôle précis. (Par des factorisations basées sur les propriétés statistiques connues des images) sur le compromis entre optimalité, au minimum erreur de Bayes, et l'efficacité de calcul [13].

Nous avons adopté la maximisation de la diversité marginale (MMD) [16] comme principe directeur pour la sélection des caractéristiques.

### 3.2. Modèles biologiquement plausibles

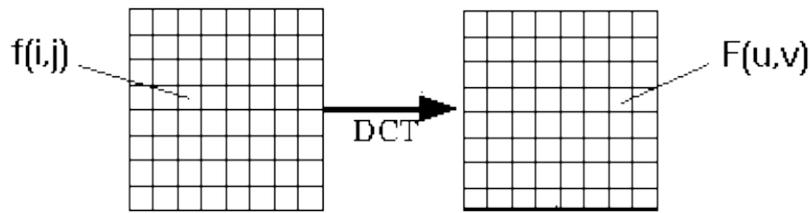
#### 3.2.1. Etape 01 : Extraction de caractéristiques (Calcul DCT)

On applique sur les caractéristiques qu'ils sont obtenir déjà après un entraînement d'apprentissage un outil mathématique de traitement du signal numérique appel Transformée en Discrète Cosinus TCD (en anglais : Discrete Cosine Transform (DCT)).

##### 3.2.1.1. Transformation en cosinus discrète (DCT)

La transformée en cosinus discrète (DCT) permet de séparer l'image en parties (ou sous-bandes spectrales) d'importance différente (en ce qui concerne la qualité visuelle de l'image).

La DCT est similaire à la transformée de Fourier discrète : elle transforme un signal ou une image du domaine spatial en domaine fréquentiel.



### ➤ Codage DCT

L'équation générale pour un DCT 1D (N données) est définie par :

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \Lambda(i) \cdot \cos \left[ \frac{\pi \cdot u}{2 \cdot N} (2i + 1) \right] f(i) \quad (1)$$

L'équation générale pour un DCT 2D (image N sur M) est définie par :

$$F(u, v) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \left(\frac{2}{M}\right)^{\frac{1}{2}} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \Lambda(i) \cdot \Lambda(j) \cdot \cos \left[ \frac{\pi \cdot u}{2 \cdot N} (2i + 1) \right] \cos \left[ \frac{\pi \cdot v}{2 \cdot M} (2j + 1) \right] \cdot f(i, j) \quad (2)$$

### ➤ Le fonctionnement de base du DCT est le suivant :

- L'image d'entrée est N par M ;
- $f(i, j)$  est l'intensité du pixel dans la rangée  $i$  et la colonne  $j$  ;
- $F(u, v)$  est le coefficient DCT de la ligne  $k1$  et de la colonne  $k2$  de la matrice DCT.
- Pour la plupart des images, une grande partie de l'énergie du signal se trouve à basses fréquences ; ceux-ci apparaissent dans le coin supérieur gauche de la DCT.
- La compression est atteinte car les valeurs inférieures droite représentent les fréquences les plus élevées et sont souvent petites - suffisamment petites pour être négligées avec une faible distorsion visible.
- L'entrée DCT est un tableau d'entiers de 8 sur 8. Ce tableau contient le niveau d'échelle de gris de chaque pixel.
- Les pixels 8 bits ont des niveaux de 0 à 255.

#### 3.2.1.2. Transformée de Fourier discrète (DFT)

Discret Fourier Transform (DFT) est un outil mathématique de traitement du signal numérique, qui est l'équivalent discret de la transformation de Fourier continue qui est utilisée pour le traitement du signal analogique. Il ne faut pas confondre avec la transformation de Fourier rapide qui est un algorithme particulier de calcul de la transformation de Fourier discrète

### 3.2.2. Etape 02 : calculer les $R_i$

En utilisant comme caractéristiques les coefficients de la transformée en cosinus discrète (DCT).Après avoir obtenu nous filtres ont cherché les réponses où chaque filtre donne deux réponses  $R_1, R_2$ .

$$R_{2k-1} = \max [-I * F_k(x, y), 0]$$

$$R_{2k} = \max [I * F_k(x, y), 0]$$

Où  $I(x, y)$  l'image d'entrée,  $R_i(x, y), i = 1, \dots, 2n$  les canaux résultant par rectification demi-onde des sorties des filtres de saillance  $F_i(x, y), i = 1, \dots, n$ .

Le calcul de  $R_1$  et  $R_2$  est frappée par le max de produit de convolution d'un signal 2D  $I(x, y)$ ,  $I$  une image avec un filtre  $F(x, y)$ .

#### 3.2.2.1. Filtrage

##### Principe

- Déplacer le filtre sur chacun des pixels ( $P = \text{pixel d'analyse}$ ) de l'image  $I$
- Remplacer le niveau en  $P$  par une valeur dépendant des niveaux de ses voisins
- Le nombre de voisins considérés dépend de la taille du filtre [15]

**2 principaux types de filtres** : le niveau en  $P$  est remplacé par

- Une combinaison linéaire des niveaux voisins (filtre linéaire, réalisable par convolution)

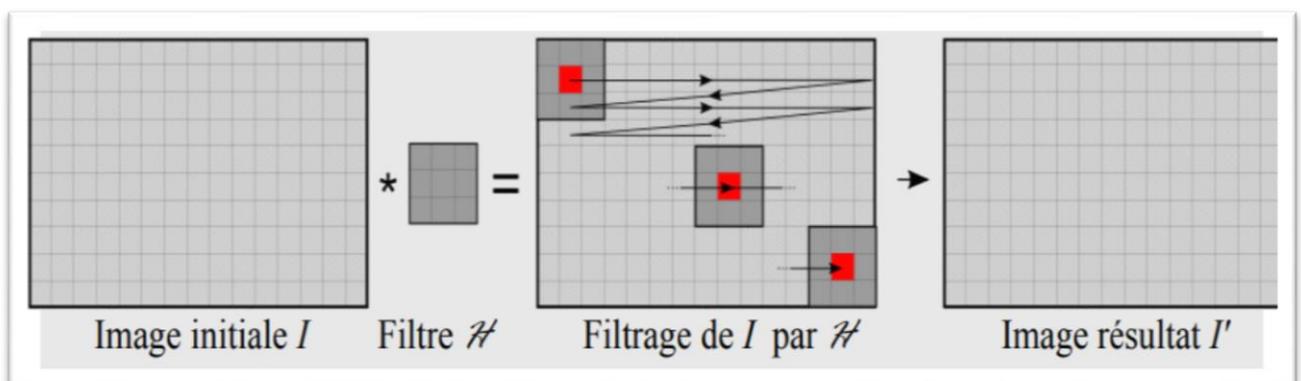


Figure 20 : Filtrage de l'image  $I$  par le filtre  $H$

- Une transformation non linéaire des niveaux voisins (filtre non linéaire) [15].

### 3.2.2.2. Filtrage par convolution discrète 2D

#### Principe :

- L'image  $I$  est une fonction de 2 variables discrètes ( $x$  et  $y$ ).
- Le filtre de convolution  $H$  appliqué sur  $I$  est lui aussi à 2D (matrice).
- $H$  est appelé filtre, masque, noyau ou fenêtre de convolution.
  - Souvent carré et de taille impaire (3x3, 5x5, ...), pour être centré sans ambiguïté sur le pixel d'analyse.
  - Souvent à valeurs symétriques par rapport à l'élément central :  $h_{-1,-1} = h_{+1,+1}$ ,  $h_{-1,0} = h_{+1,0}$ ,  $h_{0,-1} = h_{0,+1}$ , ...
  - Souvent à somme unité (normalisé) pour conserver la luminance de l'image.

$$H = \begin{bmatrix} h_{-1,-1} & h_{0,-1} & h_{+1,-1} \\ h_{-1,0} & h_{0,0} & h_{+1,0} \\ h_{-1,+1} & h_{0,+1} & h_{+1,+1} \end{bmatrix}$$

Figure 21 : Filtre de convolution

#### Formule :

- L'image  $I'$  résultat de la convolution (notée  $*$ ) de  $I$  par  $H$  est donnée par :

$$I'(x, y) = (I * H)(x, y) = \sum_{n=1}^N \sum_{m=1}^M I(x, y) H(n - x, m - y)$$

- C'est un filtrage linéaire (somme de multiplications entre les niveaux de gris de l'image initiale  $I$  et les coefficients du masque  $H$ ) [15].

#### Calcul pratique :

- Tourner le masque de 180° autour de son centre (0,0) ;
- Superposer le masque obtenu à l'image  $I$  de sorte qu'il soit centré en  $(x,y)$  ;
- Multiplier chaque coefficient du masque par le niveau du pixel sous-jacent ;
- Additionner chacun de ces produits pour obtenir  $I'(x,y)$  [15].

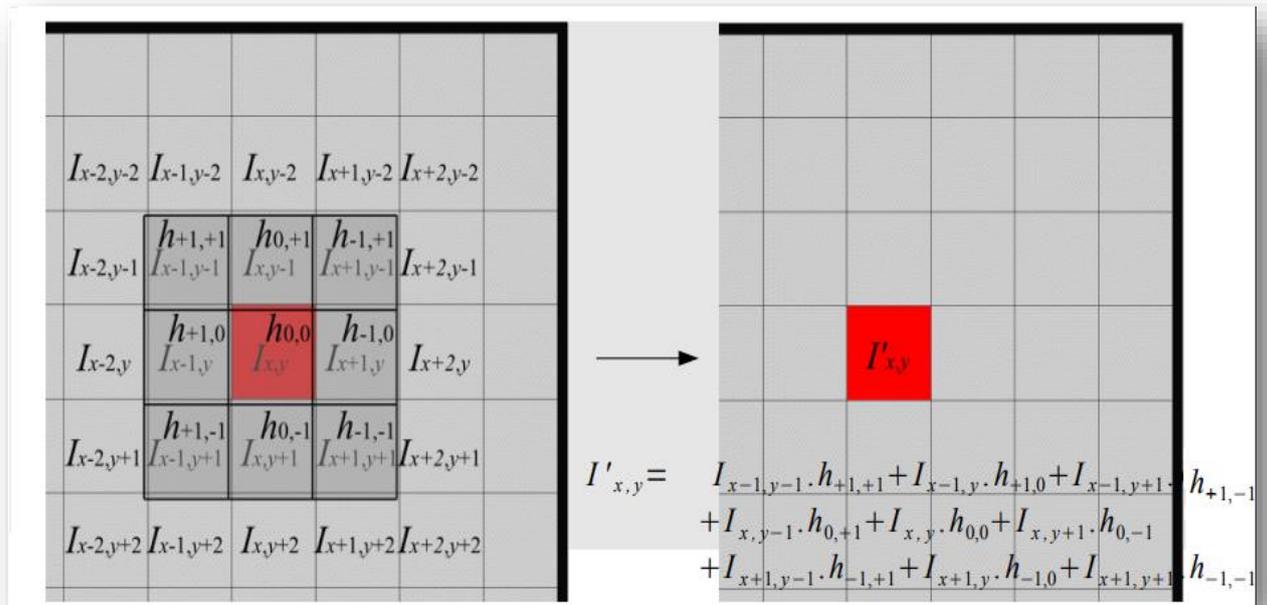


Figure 22 : Filtrage par convolution discrète 2D

On voit que la relation ne peut s'appliquer sur les bords de l'image, plus précisément sur les P rangées horizontales et verticales des bords. Il faut donc prévoir un traitement spécial pour ces rangées. La solution la plus simple, consiste à remplir ces rangées avec un niveau constant (par exemple noir). Une autre solution est de répliquer sur ces rangées les rangées voisines. Dans ces deux cas, l'information sur les bords est perdue. On peut aussi décider de ne pas modifier ces rangées, ce que nous allons faire. En tout cas, on évite de réduire la taille de l'image [15].

### 3.2.3. Etape 03 : calculer les poids $w_i$

#### 3.2.3.1. Maximum marginal diversity

Considérons un problème de classification sur un espace de caractéristiques  $X$  et un vecteur aléatoire  $X = (X_1, \dots, X_n)$  à partir duquel des vecteurs de caractéristiques sont dessinés. Alors  $md(X_k) = \langle KL [P_{x_k} | Y(x | i) || P_{x_k}(x)] \rangle_y$  est la diversité marginale de la caractéristique  $X_k$ .

L'intuition véhiculée par l'exemple ci-dessus peut être facilement transformée en un principe générique pour la sélection de caractéristiques [16].

#### Principe :

La meilleure solution pour un problème de sélection de caractéristiques consiste à sélectionner le sous-ensemble de caractéristiques conduisant à un ensemble de densités marginales extrêmement diverses [16].

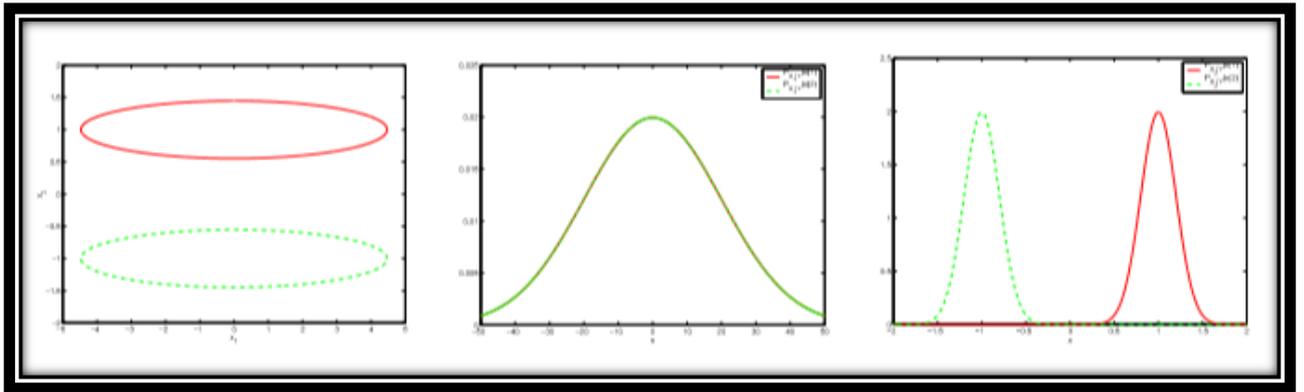


Figure 23 : Problème gaussien avec deux classes  $Y=\{1, 2\}$ , dans les deux dimensions,  $X = (X1, X2)$ .

Gauche: contours de probabilité de 65%. Milieu: marginaux pour  $X1$ . Droite: marginaux pour  $X2$ .

Ce principe a deux propriétés intéressantes. Premièrement, il est intrinsèquement discriminant, recommandant l'élimination des dimensions le long desquelles les projections des densités de classe sont les plus similaires. Deuxièmement, il est simple à implémenter avec l'algorithme suivant [16].

**Algorithme** :(sélection de caractéristiques MMD) Pour un problème de classification avec  $n$  caractéristiques  $X= \{X_1, \dots, X_n\}$ ,  $M$  classes  $Y= \{1, \dots, M\}$  et a priori de classe  $P_y(i) = P_i$  la procédure suivante renvoie les  $N$  principales fonctionnalités de MMD.

- Foreach caractéristique  $k \in \{1, \dots, N\}$ 
  - Foreach classe  $i \in \{1, \dots, M\}$  calculer une estimation d'histogramme  $h_{k,i}$  de  $P_x$
  - Calculer  $h_k = \frac{1}{M} \sum_i h_{k,i}$
  - calculer la diversité marginale  $md(X_k) = \sum_i P_i h_{k,i}^T \log(h_{k,i} / h_k)$  où **log** et la division  $/$  sont exécutés élément par élément [16].

### 3.2.3.2. Histogramme

Un histogramme un spectre de répartition d'intensité. Concrètement, il s'agit d'une liste contenant pour chaque valeur possible de pixel le nombre de pixels ayant cette valeur. Vous pouvez bien dire, mais au final, comment cela peut-il être vraiment utile? à quelles fins? Par exemple, calculez l'histogramme d'une zone sur une photographie, puis la même zone sur une autre photographie d'un autre point de vue, par exemple, aura probablement le même histogramme. Il peut donc être très utile de faire la reconnaissance de formes «à partir de zéro». En effet, OpenCV fournit toutes les fonctions nécessaires pour le faire, grâce à CalcHist et à CompareHist qui renvoient un flottant de la distance (précision) [14].

L'intention de la fonction modifie une image pour aplatir l'histogramme de l'image de sorte que la balance entre le noir et le blanc soit presque identique. L'exemple ci-dessous calcule l'histogramme d'une image grise et affiche avec un graphique l'histogramme associé (sur une autre image) [14].

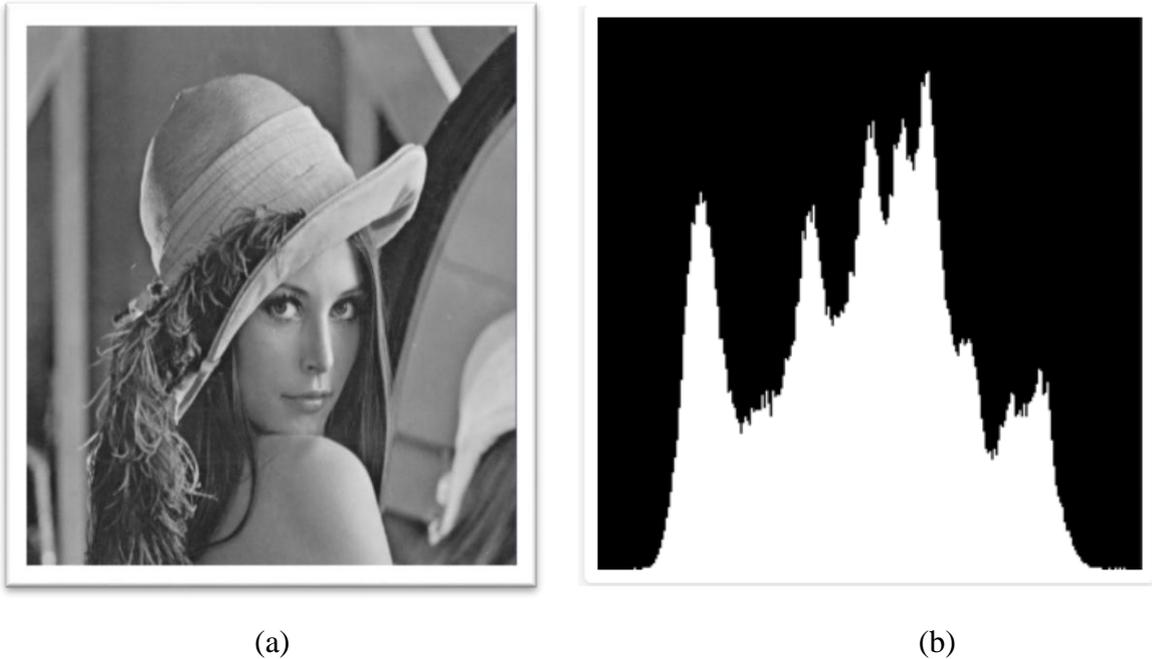
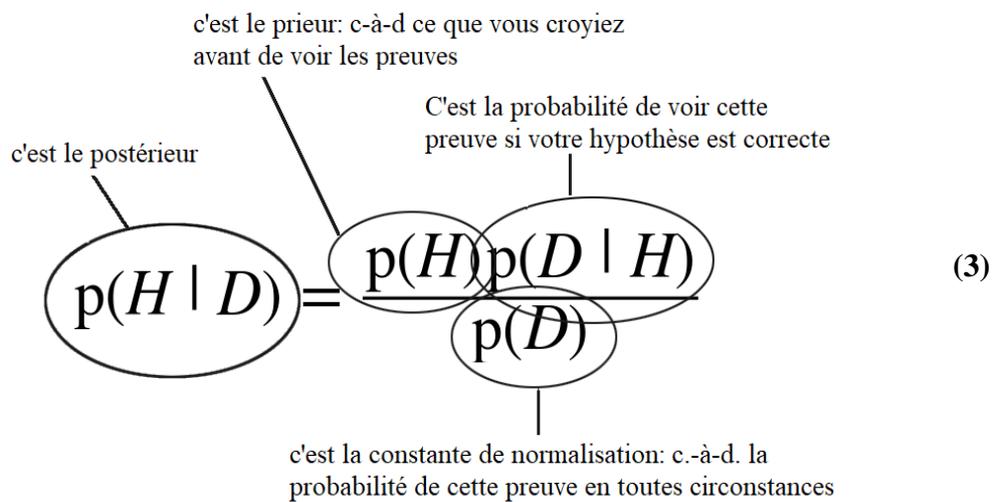


Figure 24 : (a) L'image originale, (b) Histogramme de l'image

### 3.2.3.3. Théorème de Bayes naïf



« Classification Naïve Bayes » est une technique d'apprentissage automatique supervisée. C'est une technique simple mais l'une des techniques de classification les plus efficaces. Certaines hypothèses sont formulées dans les Bayes naïves. Même si ces hypothèses sont un peu violées, cela fonctionne quand même assez bien. Hypothèses faites dans Naïve Bayes :

1- Tous les échantillons sont indépendantes, identiquement distribués, c'est-à-dire que toutes les variables aléatoires sont indépendantes les unes des autres et sont tirées de la même distribution.

2- Toutes les fonctionnalités sont conditionnellement indépendantes.

Intéressons-nous au théorème sur lequel est construit le classificateur de Naïve Bayes, c'est-à-dire le théorème de Bayes. Le théorème de Bayes stipule que [17].

$$\text{probabilité postérieure} = \frac{\text{probabilité conditionnelle} \cdot \text{probabilité préalable}}{\text{Preuve}}$$

### 3.2.3.3.1. Probabilité postérieure

La probabilité postérieure, dans le contexte d'un problème de classification, peut être interprétée comme suit : «Quelle est la probabilité qu'un objet particulier appartienne à la classe  $i$  compte tenu de ses valeurs de caractéristique observées?

Par exemple,

$$\text{Ici,} \quad P(c_j | x_i) = \frac{P(x_i | c_j) \cdot P(c_j)}{P(x_i)} \quad (4)$$

$c_j$  représente la  $j$ me classe de classes  $\{1,2,3 \dots n\}$

$x_i$  représente le vecteur de caractéristiques de son échantillon d'échantillons  $\{1,2,3 \dots m\}$

La probabilité postérieure signifie simplement : "étant donné le vecteur caractéristique  $x_i$ , quelle est la probabilité qu'un échantillon appartienne à la classe  $c_j$ ?"

### Fonction objectif de Naive Bayes

Maximiser la probabilité postérieure compte tenu des données d'apprentissage pour formuler une règle de décision pour les nouvelles données. La règle de décision pour le problème ci-dessus peut être formulée comme suit :

Si un échantillon appartient à la classe  $j$ , alors  $P(c_j | x_i)$  sera maximal. [17]

En d'autres termes, pour chaque exemple de test  $i$ , nous trouvons :

$$\arg \max_i P(c_j | x_i), \quad \forall j \in \{1, \dots, n\} \quad (5)$$

### 3.2.3.3.2. Probabilité de classe conditionnelle

Selon l'hypothèse naïve faite ci-dessus, les probabilités conditionnelles de classe ou (vraisemblance) des échantillons peuvent être directement estimées à partir des données d'apprentissage au lieu d'évaluer toutes les possibilités de  $x$ . Ainsi, étant donné un vecteur de caractéristiques  $d$ -dimensionnel  $x$ , la probabilité conditionnelle de classe peut être calculée comme suit :

$$P(x | c) = P(x_1 | c_j) \cdot P(x_2 | c_j) \cdot \dots \cdot P(x_d | c_j)$$

$$= \prod_{k=1}^d P(x_k | c_j)$$

Ici,  $P(x|c_j)$  signifie simplement : «Quelle est la probabilité d'observer ce modèle particulier  $x$  étant donné qu'il appartient à la classe  $c_j$ ?». Les probabilités «individuelles» de chaque entité du vecteur de caractéristiques peuvent être estimées via le paramètre maximum-estimation de vraisemblance, qui est simplement une fréquence dans le cas de données catégoriques :

$$\text{Où, } \hat{P}(x_i | c_j) = \frac{N_{x_i, c_j}}{N_{c_j}} \quad \forall i = 1, 2, \dots, d \quad (6)$$

$N_{x_i, c_j}$ : Nombre de fois où la fonction  $x_i$  apparaît dans les échantillons de la classe  $c_j$ .

$N_{c_j}$ : nombre total de toutes les fonctions de la classe  $c_j$ . [17]

### 3.2.3.3.3. Probabilité préalable

Dans le contexte de la classification par modèle, les probabilités antérieures sont également appelées classes antérieures de classe, qui décrivent «la probabilité générale de rencontrer une classe particulière». Ici,

$P(c_j)$  = Probabilité qu'un nouveau document appartienne à la classe  $j$ .

Si le préalable à une distribution uniforme, alors elle sera la même pour toutes les classes et la probabilité postérieure ne dépendra que de l'évidence et de la probabilité conditionnelle de classe. Comme les preuves sont communes à toutes les classes, la probabilité postérieure ne dépendra que de la probabilité conditionnelle de classe. [17]

Des connaissances préalables peuvent être obtenues auprès des experts ou estimées à l'aide des données de formation disponibles. Pour la méthode ultérieure, il est nécessaire que les données d'apprentissage soient indépendantes, identiquement distribués. et échantillon

représentatif de l'ensemble de la population. L'estimation du maximum de vraisemblance pour le précédent :

$$P(x) = \frac{Nc_j}{Nc} \quad (7)$$

$Nc_j$ : Nombre d'échantillons de la classe  $c_j$ .

$Nc$ : Nombre total d'échantillons.

#### 3.2.3.3.4. Preuve

La preuve  $P(x)$  peut être comprise comme la probabilité de rencontrer un motif particulier  $x$  indépendant du libellé de classe.

$$P(c_j | x_i) = \frac{P(x_i | c_j) \cdot P(c_j)}{P(x_i)} \quad (8)$$

Nous savons que,

$$P(c_j | x_i) + P(\neg c_j | x_i) = 1$$

Alors,

$$P(x) = P(x_i | c_j) \cdot P(c_j) + P(x_i | \neg c_j) \cdot P(\neg c_j)$$

Bien que le terme de preuve soit nécessaire pour calculer avec précision les probabilités postérieures, il peut être supprimé de la règle de décision, car sa seule contribution à la classification est la normalisation. [17]

### 3.2.3.3.5. Lissage additif

Si, au cours de la période de test, nous rencontrons une caractéristique que nous n'avons pas rencontrée pendant la période d'entraînement, la probabilité conditionnelle individuelle de cette caractéristique devient nulle, ce qui donne une probabilité de condition conditionnelle de classe égale à zéro. Nous devons donc modifier la formule de calcul de la probabilité conditionnelle individuelle. Afin d'éviter le problème des probabilités nulles, un terme de lissage supplémentaire peut être ajouté au modèle multinomial de Bayes. Les variantes les plus courantes du lissage additif sont le lissage dit de Lidstone ( $\alpha < 1$ ) et le lissage de Laplace ( $\alpha = 1$ ). [17]

$$\hat{P}(x_i | c_j) = \frac{N_{x_i, c_j} + \alpha}{N_{c_j} + \alpha d} \quad \forall i = 1, 2, \dots, d \quad (9)$$

$N_{x_i, c_j}$ : Nombre de fois où la fonction  $x_i$  apparaît dans les échantillons de la classe  $c_j$ .

$N_{c_j}$ : nombre total de toutes les fonctions de la classe  $c_j$ .

$\alpha$ : paramètre de lissage additif

$d$ : dimensionnalité du vecteur de caractéristiques  $x$

### 3.2.4. Etape 04 : Calcule de la saillance (la somme)

-Une carte de saillance est obtenue en regroupant les réponses des différents filtres de saillance  $R_i$  multiplié avec les poids  $W_i$

$$S(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{2n} w_i \cdot R_i^2(\mathbf{x}, \mathbf{y}) \quad (10)$$

## 4. Schéma global :

Nous avons défini, jusqu'à présent les éléments principaux permettant de concevoir une technique de création d'une carte de saillance en utilisant la théorie décisionnelle, se basant sur la notion de classes et de caractéristiques liées à la DCT.

Nous allons fournir ci-dessous, le schéma global de la technique :

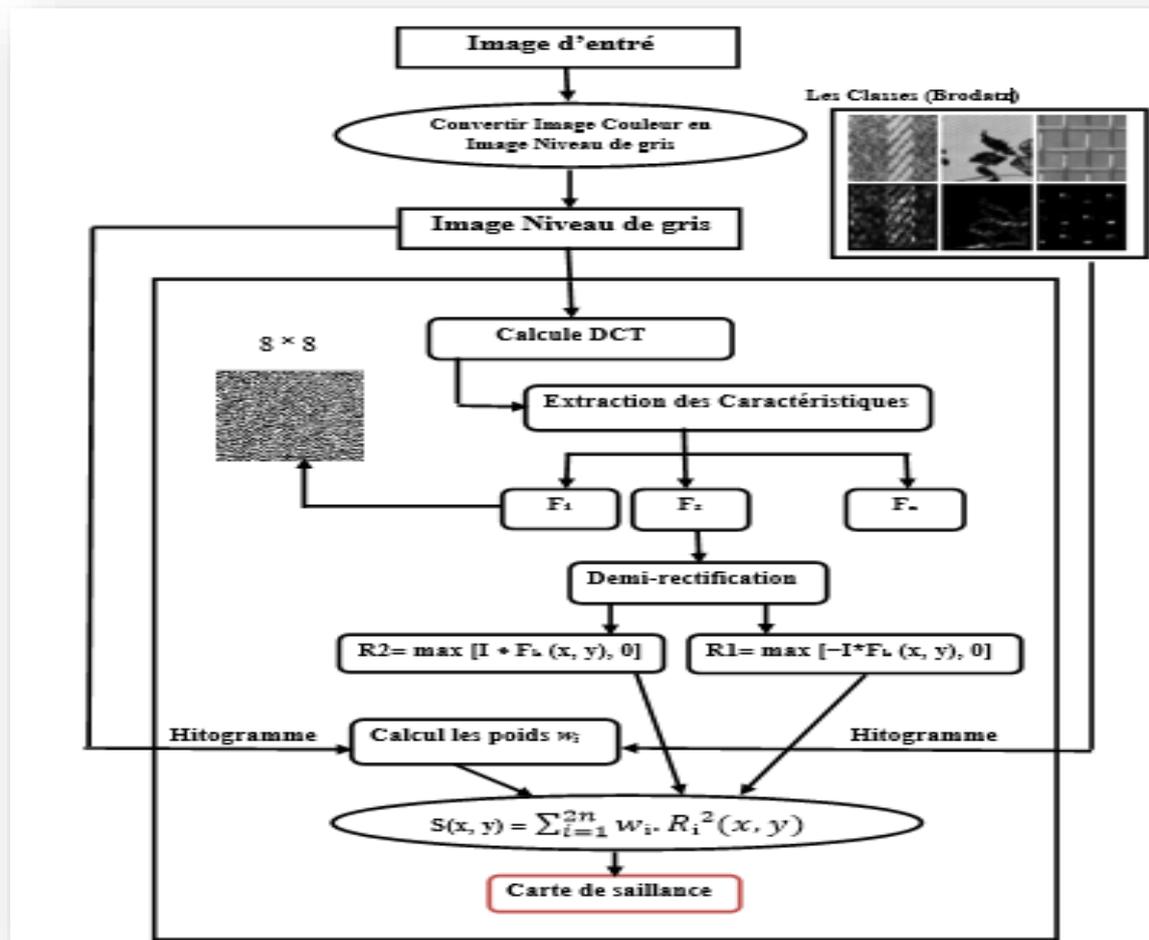


Figure 25 : Schéma globale de modèle de Discriminante

## 5. Conclusion

Dans ce chapitre nous avons expliqué chaque étape du modèle de perception visuelle «Modèle de théorie de décision » pour extraire les caractéristiques visuelles dans le but de construire une carte de saillance. Nous parlons sur l'implémentation de ce modèle dans le chapitre suivant.

# **Chapitre 04 : Implémentation, résultats et bilan**

## 1. Introduction

Après avoir élaboré une étude détaillée sur le principe du modèle de l'attention visuelle basée sur la théorie de décision. Nous consacrons le chapitre courant aux aspects techniques de notre travail, afin d'implémenter la méthode de discrimination. Nous nous intéressons donc aux résultats obtenus après la réalisation de notre application et pour les validés, nous les comparons aux résultats obtenus par la méthode de base de Itti. Une deuxième validation a été faite en changeant le nombre de caractéristiques pour voir la variation des résultats. D'abord, nous allons présenter les langages de programmation et les outils de développement utilisés lors de l'implémentation de notre application. Par la suite, nous décrivons l'application réalisée.

## 2. Outils de développement

Microsoft Visual Studio est une suite de logiciels de développement pour Windows et mac OS conçue par Microsoft. La dernière version s'appelle Visual Studio 2019.

Visual Studio est un ensemble complet d'outils de développement permettant de générer des applications web ASP.NET, des services web XML, des applications bureautiques et des applications mobiles. Visual Basic, Visual C++, Visual C# utilisent tous le même environnement de développement intégré (IDE), qui leur permet de partager des outils et facilite la création de solutions faisant appel à plusieurs langages. Par ailleurs, ces langages permettent de mieux tirer parti des fonctionnalités du framework .NET, qui fournit un accès à des technologies clés simplifiant le développement d'applications web ASP et de services web XML grâce à Visual Web Developer.



*Figure 26 : Logo de Visual Studio 2010*

### 2.1.Langages de programmation

C++ est un langage de programmation compilé permettant la programmation sous de multiples paradigmes (comme la programmation procédurale, orientée objet ou générique). Ses bonnes performances, et sa compatibilité avec le C en font un des langages de programmation les plus utilisés dans les applications où la performance est critique.

Le langage de programmation Visual C ++ et les outils de développement aident à développer des applications Windows Universal natives, des applications de bureau et de serveur natives, des bibliothèques multiplateformes fonctionnant sous Android et iOS, ainsi que Windows et des applications gérées sur .NET Framework.

## 2.2. Bibliothèques utilisées

OpenCV (Open Source Computer Vision Library) est une bibliothèque de logiciels open source de vision informatique et d'apprentissage automatique. OpenCV a été conçu pour fournir une infrastructure commune aux applications de vision par ordinateur et pour accélérer l'utilisation de la perception de la machine dans les produits commerciaux. OpenCV étant un produit sous licence BSD, il est facile pour les entreprises d'utiliser et de modifier le code.

La bibliothèque contient plus de 2500 algorithmes optimisés, qui incluent un ensemble complet d'algorithmes classiques et avancés de vision par ordinateur et d'apprentissage automatique. Ces algorithmes peuvent être utilisés pour détecter et reconnaître des visages, identifier des objets, classer les actions humaines dans des vidéos, suivre les mouvements d'une caméra, suivre des objets en mouvement, extraire des modèles 3D d'objets, produire des nuages de points 3D à partir de caméras stéréo, assembler des images pour obtenir une haute résolution. l'image d'une scène entière, trouver des images similaires dans une base de données d'images, supprimer les yeux rouges des images prises au flash, suivre les mouvements des yeux, reconnaître les paysages et établir des repères pour les superposer à la réalité augmentée, etc. OpenCV compte plus de 47 000 utilisateurs communauté et le nombre de téléchargements estimé à plus de 18 millions. La bibliothèque est largement utilisée dans les entreprises, les groupes de recherche et les organismes gouvernementaux.

Il possède des interfaces C ++, Python, Java et MATLAB et prend en charge Windows, Linux, Android et Mac OS. OpenCV s'appuie principalement sur les applications de vision en temps réel et tire parti des instructions MMX et SSE, le cas échéant. Des interfaces CUDA et OpenCL complètes sont activement développées en ce moment. Il existe plus de 500 algorithmes et environ 10 fois plus de fonctions qui composent ou supportent ces algorithmes. OpenCV est écrit nativement en C ++ et possède une interface basée sur un modèle qui fonctionne de manière transparente avec les conteneurs STL.

### 3. Implémentation

#### 3.1. Diagramme de classe de modèle conceptuel

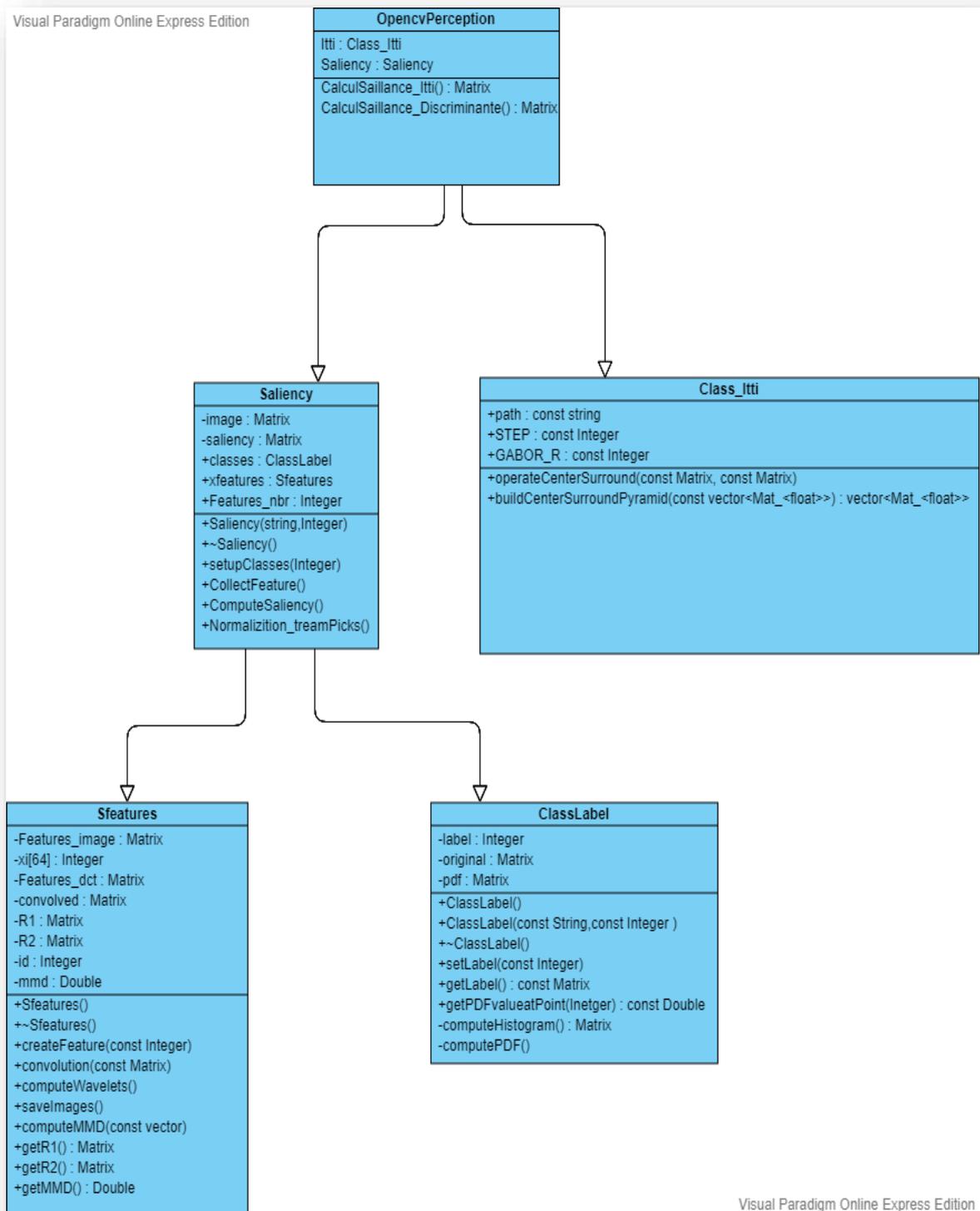


Figure 27 : Diagramme de Classe de modèle implémenter

La figure suivante (figure 28) représente l'interface de notre application qui affiche l'image originale dans la zone (a), (b) la carte de saillance de modèle Itti et (c) la carte de saillance de modèle de la théorie de décision (carte de saillance de modèle discriminante).

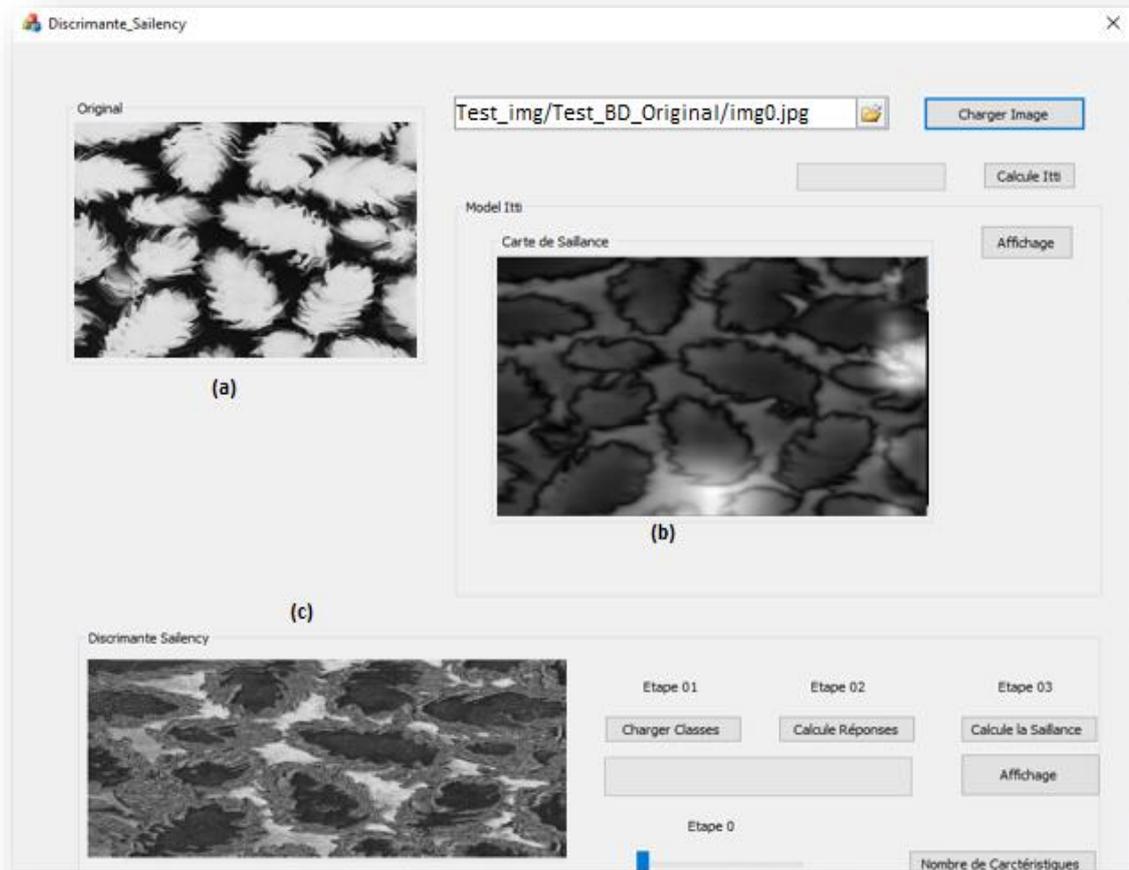


Figure 28 : Capture d'écran présente notre application en mode de fonction qui affiche les images (a) image original, (b) carte de saillance Itti, (c) carte de saillance du modèle discriminante.

### 3.2.Procédures et Fonctions utilisées

Dans cette section nous allons présenter quelques fonctions que nous avons programmées pour implémenter l'application finale.

#### *Etape 01 : Extraction de caractéristiques (Calculer DCT)*

```
void Sfeatures::createFeature(const int& nbr)
{
    char* imagefileName = new char[255];
    sprintf(imagefileName, "dataset/DCTbasis/base%d.jpg", nbr);
    std::string buffer(imagefileName);

    this->Features_image = cv::imread(buffer, CV_LOAD_IMAGE_GRAYSCALE);

    if(!this->Features_image.data)
    {
        throw std::invalid_argument("no image!!");
    }
    //calculer DCT

    cv::Mat complexImg = cv::Mat(8, 8, CV_32FC1, cv::Scalar(0.0));
    this->Features_image.convertTo(complexImg, CV_32F);
    cv::Mat temp;
    cv::dct(complexImg, temp);
    // save dct
    this->Features_dct = cv::Mat(8, 8, CV_32FC1, cv::Scalar(0.0));
    int counter = 0;
    for(int y = 0; y < 8; y++)
    {
        for(int x = 0; x < 8; x++)
        {
            this->Features_dct.at<float>(y,x) = temp.at<float>(y,x);
            this->xi[counter] = (int)this->Features_image.at<uchar>(y,x); // convert int (getPDF)
            counter++;
        }
    }

    this->id = nbr;
    this->mmd = 0.0; // wi

    std::cout<<"feature created"<<std::endl;
}
```

Figure 29 : Capture d'écran présente la procédure qui charge les caractéristiques et calcule la DCT

**Étape 02 : Calcul des  $R_i$** 

```

void Sfeatures::convolution(const cv::Mat& origine)
{
    this->convolved = cv::Mat(origine.rows, origine.cols, CV_8SC1, cv::Scalar(0));
    //cv::Mat flipped;
    //cv::flip(Features_dct, flipped, -1);
    //cv::filter2D(origine, convolved, -1, Features_dct);

    for(int y = 0; y < origine.rows; y++)
    {
        for(int x = 0; x < origine.cols; x++)
        {
            float sum = 0.0;
            for (int k=-4 ; k<=3 ; k++)
            {
                for (int j=-4 ; j<=3 ; j++)
                {
                    int x1 = reflect(origine.cols, x - j);
                    int y1 = reflect(origine.rows, y - k);
                    sum += (Features_dct.at<float>(k+4,j+4) * (float)origine.at<uchar>(y1,x1));
                }
            }
            this->convolved.at<char>(y,x) = (char)sum;
        }
    }

    std::cout<<"feature convolved"<<std::endl;
}

```

Figure 30 : Capture d'écran présente la procédure de filtrage a convolution

```

void Sfeatures::computeWavelets()
{
    this->R1 = cv::Mat(this->convolved.rows, this->convolved.cols, CV_8SC1);
    this->R2 = cv::Mat(this->convolved.rows, this->convolved.cols, CV_8SC1);

    for(int y = 0; y < this->R1.rows; y++)
    {
        for(int x = 0; x < this->R1.cols; x++)
        {
            float val = (float)this->convolved.at<char>(y,x);
            this->R1.at<char>(y,x) = (char)std::max(val,0.0f);
            this->R2.at<char>(y,x) = (char)std::max(-val,0.0f);
        }
    }
    std::cout<<"wavelet computed"<<std::endl;
}

```

Figure 31 : Capture d'écran présente la procédure de Calcule les Réponses ( $R_i$ )

**Etape 03 : calculer les poids  $w_i$** 

```

void Sfeatures::computeMMD(const std::vector<ClassLabel> &labes)
{
    std::vector<cv::Mat> H_k_i;
    for(std::size_t i = 0; i < labes.size(); ++i)
    {
        cv::Mat hki( 1, 64, CV_64FC1, cv::Scalar(0.0) );
        for(int d = 0; d < 64; ++d)
            hki.at<double>(d) = labes[i].getPDFvalueatPoint(this->xi[d]);
        H_k_i.push_back(hki);
    }
    std::vector<cv::Mat> H_k;
    for(std::size_t i = 0; i < labes.size(); ++i)
    {
        cv::Mat hk( 1, 64, CV_64FC1, cv::Scalar(0.0) );
        for(int d = 0; d < 64; ++d)
        {
            for(std::size_t j = 0; j < labes.size(); ++j)
                hk.at<double>(d) += H_k_i[j].at<double>(d);
        }
        for(int d = 0; d < 64; ++d)
        {
            hk.at<double>(d) /= (double)labes.size();
        }
        H_k.push_back(hk);
    }
    //compute log
    std::vector<cv::Mat> logH;
    for(std::size_t i = 0; i < labes.size(); ++i)
    {
        cv::Mat t( 64, 1, CV_64FC1, cv::Scalar(0.0) );
        for(int d = 0; d < 64; ++d)
        {
            for(std::size_t j = 0; j < labes.size(); ++j)
                t.at<double>(d) = log(H_k_i[j].at<double>(d) / H_k[j].at<double>(d));
        }
        logH.push_back(t);
    }
    double p_i = 1.0 / labes.size(); // class prior
    for(std::size_t i = 0; i < labes.size(); ++i)
    {
        cv::Mat x = ((p_i * H_k_i[i]) * logH[i]);
        this->mmd += x.at<double>(0);
    }
}

```

*Figure 32 : Capture d'écran présente la procédure de Calcule MMD*

```
cv::Mat ClassLabel::computeHistogram()
{
    cv::Mat histogram;
    int histSize = 256;

    float range[] = { 0, 256 };
    const float* histRange = { range };

    bool uniform = true; bool accumulate = false;
    cv::calcHist( &this->original, 1, 0, cv::Mat(), histogram, 1, &histSize, &histRange, uniform, accumulate );

    return histogram;
}
```

Figure 33 : Capture d'écran présente la fonction de calcul de l'histogramme

```
void ClassLabel::computePDF()
{
    cv::Mat hist;
    hist = computeHistogram();

    double sum = 0.0;
    cv::Mat temp( 1, 257, CV_64FC1, cv::Scalar(0.0) );

    for( int i = 1; i < 257; i++ )
        sum += hist.at<float>(i-1);

    // apply Additive Smoothing
    for( int i = 1; i < 257; i++ )
        temp.at<double>(i-1) = (hist.at<float>(i-1) + 1.0) / (sum + 64.0);

    this->pdf = temp;
}
```

Figure 34 : Capture d'écran présente la procédure de calcul de PDF

## Étape 04 : Calcul de la saillance (la somme)

```

void Saliency::ComputeSaliency()
{
    this->saliency = cv::Mat(this->image.rows, this->image.cols, CV_64FC1, cv::Scalar(0.0));

    for(int y = 0; y < this->saliency.rows; y++)
    {
        for(int x = 0; x < this->saliency.cols; x++)
        {
            for(std::size_t i = 0; i < this->xfeatures.size(); ++i)
            {
                cv::Mat r1 = this->xfeatures[i].getR1();
                cv::Mat r2 = this->xfeatures[i].getR2();
                double wi = this->xfeatures[i].getMMD();

                double r1_square = (double)r1.at<char>(y,x) * (double)r1.at<char>(y,x);
                double r2_square = (double)r2.at<char>(y,x) * (double)r2.at<char>(y,x);

                this->saliency.at<double>(y,x) += (wi * (r1_square + r2_square));
            }
        }
    }

    this->Normalization_treamPicks();
    cv::normalize(this->saliency, this->saliency, 0, 255, cv::NORM_MINMAX, CV_8UC1);
    cv::imwrite( "sailency.jpg", this->saliency);
}

```

Figure 35 : Capture d'écran présente la procédure de calcul de la saillance

## 4. Tests et Comparaison

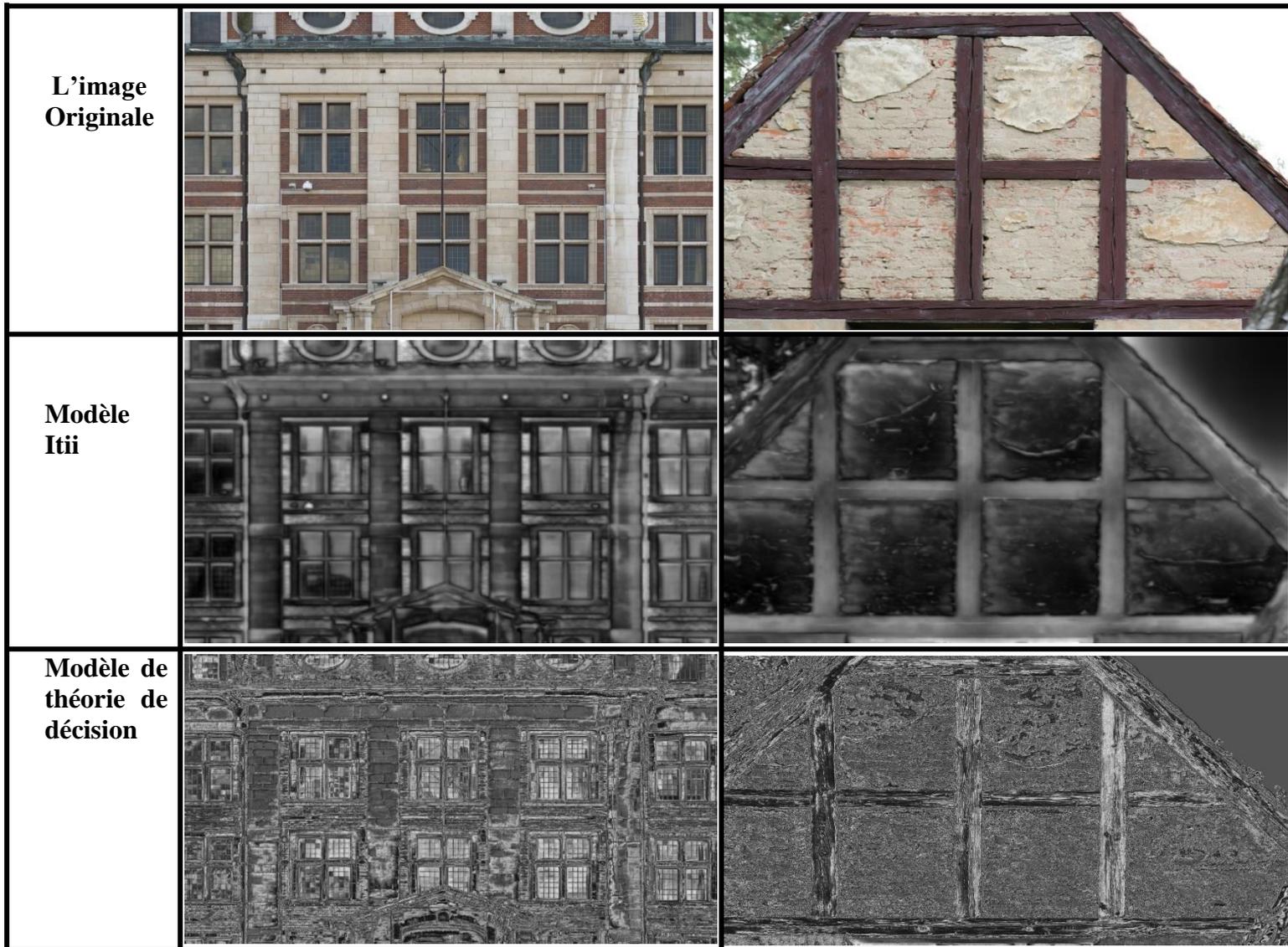
Dans cette section nous avons faire des tests sur résultats visuelle avec des déférents images de plusieurs base de données (Images du texture, Images de la base d'entraînement, Images d'Art) et de comparaisons avec le modèle de base Itti

Et pour une deuxième validation faire une Test de Résultats par rapport le nombre de Caractéristiques

### 4.1. Test des résultats Visuels

Les tests ont été élaborés sur plusieurs types d'images, nous avons essayé de faire une interprétation visuelle des résultats, mais une validation rigoureuse exige l'utilisation de métrique perceptuelle telle que VDP (Visible Difference Metric).

## 4.1.1. Base de test : Images de Texture

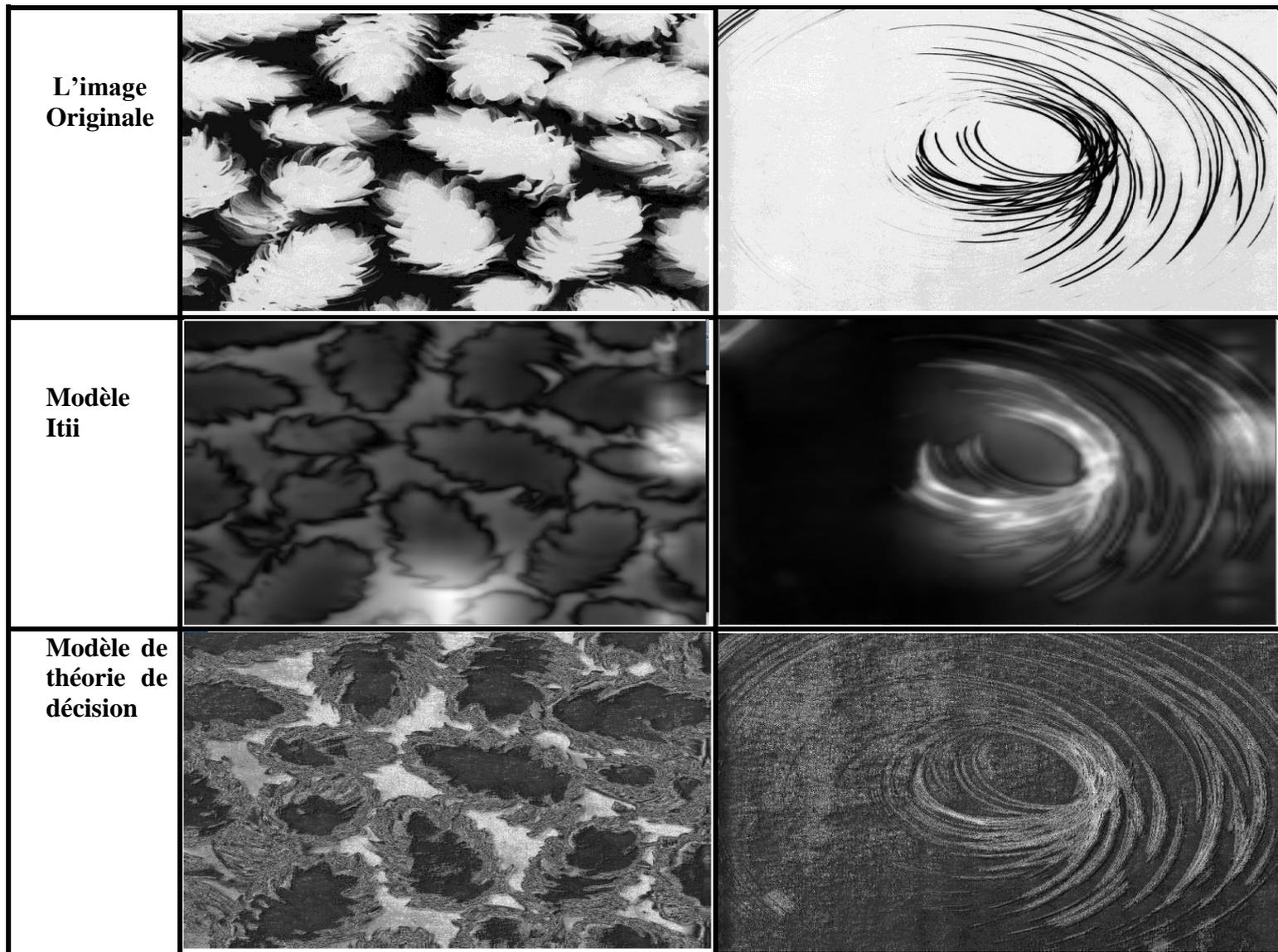


*Table 01 : Table de comparaison les résultats de modèle Itti avec le modèle de discriminante*

*(Base de données : images de texture)*

Pour la première table quand on utilise des images de texture dans l'image à gauche la carte de saillance de modèle Itti est mieux et bien contraster que la carte de saillance du modèle de discriminante, mais dans l'image à droite la carte de saillance du modèle de discriminante est contraster sur les barres verticale et horizontale qu'ils sont plus saillantes dans cette image.

## 4.1.2. Base de test : Images de la base d'entraînement

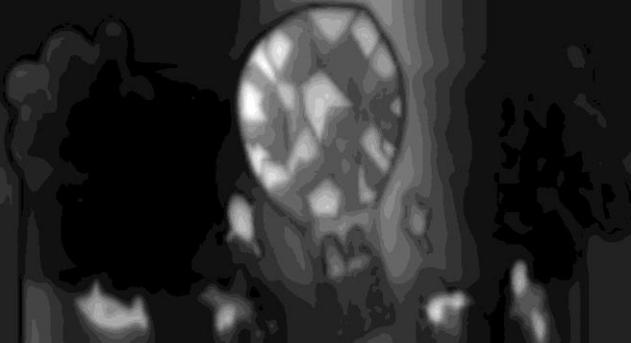


**Table 02 :** *Table de comparaison les résultats de modèle Itti avec le modèle de discriminante*

(Base de données : images de la base d'entraînement)

Pour la deuxième table quand on utilise des Images de la base d'entraînement, nous notons que la carte de saillance du modèle de discriminante est mieux défini la saillance que le modèle Itti dans les deux images. (Les parties saillantes sont plus discriminées).

## 4.1.3. Base de test : Images d'Art

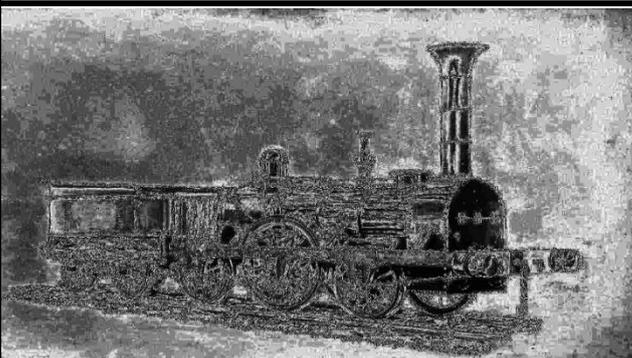
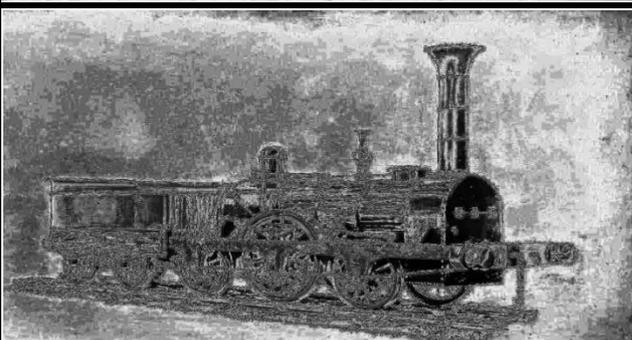
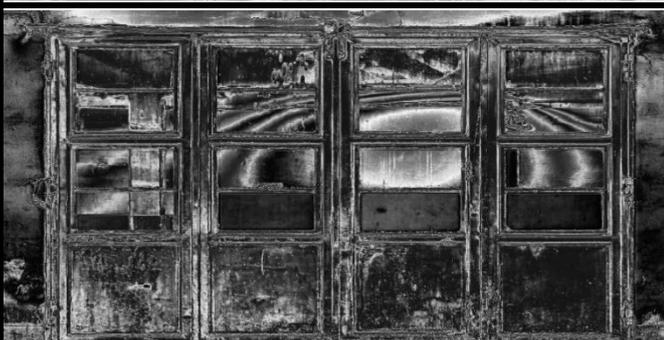
L'image Originale		
Modèle Itii		
Modèle de théorie de décision		

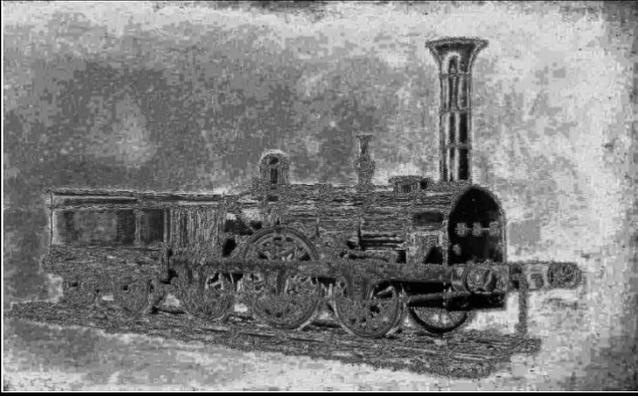
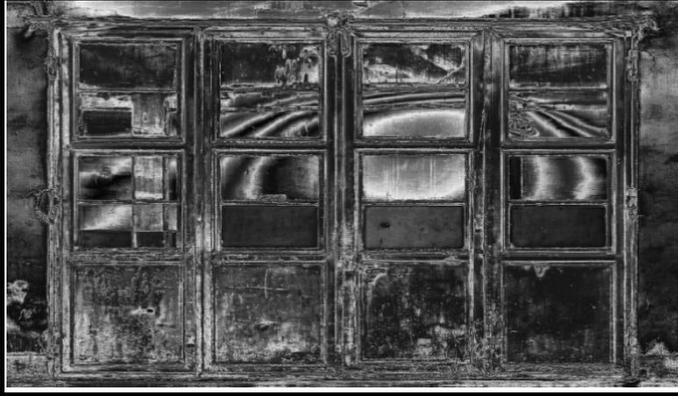
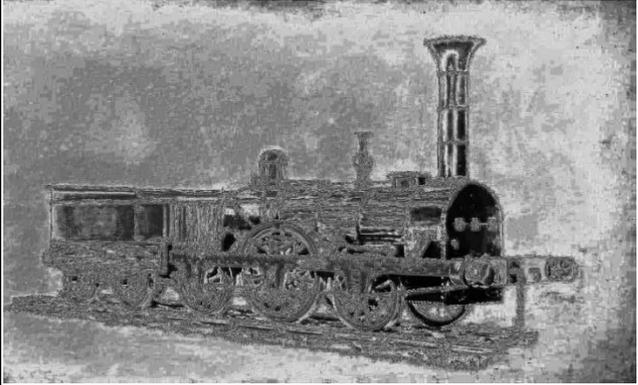
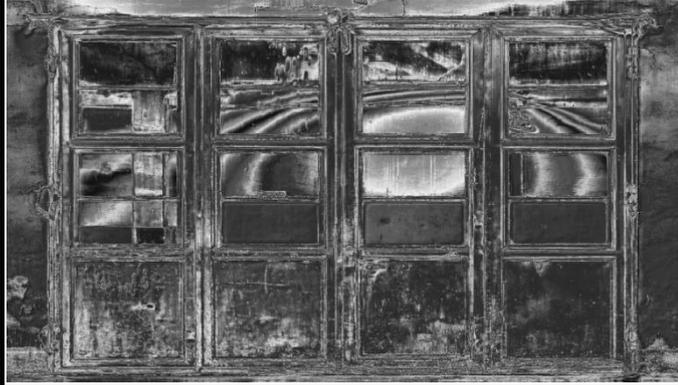
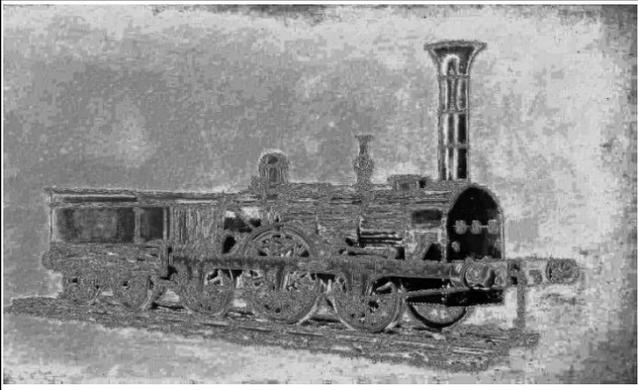
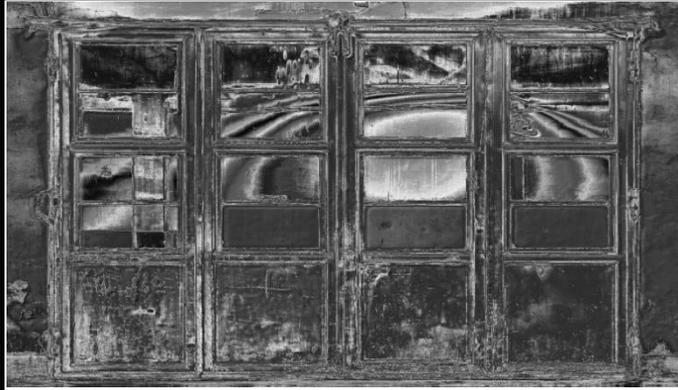
**Table 03 :** Table de comparaison les résultats de modèle Itii avec le modèle de discriminante

(Base de données : images d'Art)

Pour la troisième table, nous utilisons des Images d'Art, nous remarquons que l'image à droite la carte de saillance de modèle Itii est mieux et bien contraster que la carte de saillance de le modèle de discriminante, mais dans l'image à gauche la carte de saillance de notre modèle est plus contrastée sur les nuages qui sont plus saillants dans cette image.

4.2. Test de Résultats par rapport au nombre de Caractéristiques

<p><b>L'image Originale</b></p>		
<p><b>Modèle Itii</b></p>		
<p><b>Modèle de théorie de décision</b> <b>Nbr de caract= 5</b></p>		
<p><b>Nbr de caract= 11</b></p>		

Nbr de caract= 17		
Nbr de caract= 25		
Nbr de caract= 30		

**Table 04 :** Table de comparaison les résultats de modèle Itti avec le modèle de discriminant de nombre de caractéristique variable

Dans la table 04, nous remarquons que dans la carte de saillance la première image (image de gauche) du modèle Itti la saillance est globalement bien définie, ce qui n'est pas le cas du modèle de discriminante où il est difficile de reconnaître les parties saillantes même lorsqu'on augmente le nombre de caractéristiques.

Pour la deuxième image (image de droite), la carte de saillance du modèle de discriminante est plus significative, les objets saillants sont plus visibles et mieux cernés dans le modèle discriminant que dans la carte de saillance de modèle Itti, malgré qu'on a le même effet

même lorsqu'on augmente le nombre de caractéristique (aucun changement dans la saillance).

## **5. Conclusion**

Nous avons présenté dans ce chapitre les résultats de l'implémentation de la carte de saillance en utilisant une technique basée sur la théorie de décision ainsi qu'une validation avec les résultats obtenus par la méthode classique d'Itti.

L'objectif étant d'effectuer une première ébauche de validation, nous pouvons remarquer que la technique d'Itti qui se base sur des caractéristiques connues comme la couleur et l'orientation et la combinaison des cartes qui lui sont associées est totalement indépendante du contexte, alors que la technique basée sur la théorie de décision qui utilise principalement la DCT : Discrete Cosine Transform, qui possède une excellente propriété de « regroupement » de l'énergie : l'information est essentiellement portée par les coefficients basses fréquences, est très liée à la base d'apprentissage.

## Conclusion Générale

Dans ce travail, nous avons proposé la mise en œuvre d'un modèle de perception visuelle, c'est le modèle de théorie de décision qui permet dans une première phase Extraction de caractéristiques visuelles pertinentes, basées sur la DCT. Puis dans une seconde phase calcule les  $R_i$ , et les poids  $w_i$ , il effectue enfin une combinaison de  $w_i$ , avec le carré de chaque  $R_i$  pour créer une carte de saillance, qui affiche et sélectionne les zones importantes qui attirent l'attention visuelle humaine.

Nous avons obtenu les résultats suivants :

- les caractéristiques obtenues après un entraînement d'apprentissage sur une base données, selon le modèle de théorie de décision.
- Nous avons appliqué un outil mathématique de traitement du signal numérique : Transformée en Discrète Cosinus DCT pour chaque caractéristique. Celle-ci possède en effet une excellente propriété de « regroupement » de l'énergie : l'information est essentiellement portée par les coefficients basses fréquences.
- Puis nous avons calculé les  $R_i$  avec un filtrage à convolution par les filtres  $F_i$  (les caractéristiques),
- Ensuite, nous avons calculé les poids  $w_i$ , et nous avons appliqué l'algorithme de Maximum marginal diversity.
- Enfin, nous avons effectué une combinaison de  $w_i$ , avec le carré de chaque  $R_i$ , ce qui nous a permis d'obtenir la carte de saillance, qui est une image 2D représentant l'attention visuelle

Ce travail ouvre de nombreuses perspectives, qui peuvent être envisagées soit comme une amélioration ou une extension directe de ce travail, soit comme nécessitant des études approfondies et à plus long terme. Ces perspectives peuvent être résumées ainsi :

**La saillance pour les scènes dynamiques :** Nous pouvons envisager des perspectives nécessitant des études préalables plus approfondies. Comme par exemple étudier la saillance visuelle pour les scènes dynamiques.

**Intégration des modèles proposés dans l'application de la recherche d'image par contenu :** Plusieurs applications technologiques des modèles d'attention visuelle ont été développées au cours des années et qui ont encore augmenté l'intérêt pour la modélisation de la perception visuelle, comme les applications de vision et graphiques, robotique et d'autres applications.

**Tracking :** le suivi des objets.

## Bibliographie et webographie

- [1] Hansen, T. (2003). A neural model of early vision: Contrast, contours, corners and surfaces. University of Ulm Faculty of Computer Science, 35-42.
- [2] Kibiuk, L. V., Stuart, D., & Miller, M. (2008). *Brain Facts: A Primer on the Brain and Nervous System*. The Society For Neuroscience.
- [3] Gao, X., Lu, W., Tao, D., & Li, X. (2010, August). Image quality assessment and human visual system. In *Visual Communications and Image Processing 2010* (Vol. 7744, p. 77440Z). International Society for Optics and Photonics.
- [4] Sánchez, A. J. R. (2004). Attention, Visual Search and Object Recognition. Phd thesis, York University Toronto Canada.
- [5] Goebel, D. K. R., Mucklil, L., & Kim, D. S. (2012). The Human Nervous System, chapter Visual System. Academic Press.
- [6] M. Veyret and E. Maisel (2004) Simulation de la focalisation del'attentionvisuelle : application`a la simulation d'automobilistes virtuels, 17`emesJourn´ees de l'AFIG, November, 2004.
- [7] Itti, Laurent (2000) Models of bottom-up and top-down visual attention. Dissertation (Ph.D.), California Institute of Technology.
- [8] Itti, L. (2005). Models of Bottom-up Attention and Saliency. *Neurobiology of Attention*, 576–582. doi:10.1016/b978-012375731-9/50098-7
- [9] Thomas, M. S., & McClelland, J. L. (2008). Connectionist models of cognition. *Cambridge handbook of computational cognitive modelling*, pages 23-58.
- [10] Borji, A., Itti, L., Liu, J., Musialski, P., Wonka, P., Ye, J., ...& Yu, K. State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35(1):185-207, 2013.
- [12] Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11), 1254-1259.
- [11] ABOUELAZIZ, I., & EL HASSOUNI, M. Évaluation de la qualité visuelle des objets 3D. Projet de fin d'études. LRIT-CNRST URAC29, Université Mohammed V Rabat. 2013
- [13] Gao, D. and Vasconcelos, N. (2004). Discriminant saliency for visual recognition from cluttered scenes. *Advances in Neural Information Processing Systems*

[14] Robin David. OpenCV Tutorial Introduction, Chapter 04 Histogram and Back projection. 2018

[15] Projet Encadré IVI. filtrage d'image et compléments sur OpenCV. Université de Lille. <http://master-ivi.univ-lille1.fr/fichiers/Cours/pje-semaine-4-filtrage.pdf> (vu le 24/06/2019)

[16] N. Vasconcelos. Feature Selection by Maximum Marginal Diversity. In Neural Information Processing System, Vancouver, Canada, 2002.

[17] <https://medium.com/@akankshamalhotra24/naive-bayes-theorem-79832d506a63>  
consulté le 11/06/2019 à 12:45 .

[18] D. G. Lowe. Object recognition from local scale-invariant features. In Proceedings of International Conference on Computer Vision, pp. 1150-1157, 1999.

[19] N. Sebe, M. S. Lew. Comparing salient point detectors. Pattern Recognition Letters, vol.24, no.1-3, Jan. 2003, pp.89-96.

[20] T. Kadir and M.I Brady. Scale, Saliency and Image Description. International Journal of Computer Vision, Vol.45, No.2, p83-105, November 2001