



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 : 11SIOD/M2/2018

Mémoire

présenté pour obtenir le diplôme de master académique en

Informatique

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

L'utilisation de la perception visuelle pour l'amélioration des images

Par :

GHEBGHOUB IMANE

Soutenu le 24 juin 2018, devant le jury composé de :

ABABSA Tarek

MCB

Président

BABAHENINI Mohamed Chaouki

Professeur

Rapporteur

CHIGHOUB Fouzia

MAA

Examineur

Remerciements

Nous tenons tout d'abord à remercier Dieu le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce travail.

*En second lieu, nous tenons à remercier notre encadreur Mr le professeur **Babahenini Mohamed Chaouki**, son précieux conseil et son aide durant toute la période du travail.*

Nos vifs remerciements vont également aux membres du jury pour l'intérêt qu'ils ont porté à notre mémoire en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.

Enfin, nous tenons également à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Dédicaces

A mes chers parents, pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études.

A mes chères sœurs Chahinez, Wiam, Romaisa et à mes chers frères, AbdelMonaim, et Abderrahmen, pour leurs encouragements permanents, et leur soutien moral,

A mes chères cousines Mouna, Assma, Houda, Manar, pour leur appui et leur encouragement,

A toute ma famille, surtout mon grand-père et ma grande mère pour leur soutien tout au long de mon parcours universitaire,

Que ce travail soit l'accomplissement de vos vœux tant allégués, et le fruit de votre soutien infaillible.

Merci d'être toujours là pour moi.

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 notre mémoire, un modèle de perception visuelle, basé sur les processus Bottom-up, est présenté. Dans ce modèle, différentes caractéristiques de l'image de départ sont combinées en une carte de saillance qui est devenue de plus en plus importante avec le développement d'une technologie d'identification intelligente et de vision artificielle. Ce processus est essentiel pour de nombreuses applications de traitement d'images telles que la récupération d'image, la segmentation d'image, la reconnaissance d'image et la compression d'image adaptative, ou l'amélioration des images

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 our memoir, a visual perception model, based on Bottom-up processes, is presented. In this model, different characteristics of the starting image are combined into a saliency map that has become increasingly important with the development of intelligent identification technology and artificial vision. This process is essential for many image processing applications such as image retrieval, image segmentation, image recognition and adaptive image compression, or image enhancement

Table des matières :

Introduction générale	11
Chapitre 01 :	13
La perception visuelle et systèmes de couleurs	13
1. Introduction :	14
2. L'œil et son fonctionnement :	14
2.1. Définition :	14
2.2. Composants de l'œil :	15
3. La perception visuelle :	17
3.1. Construction de la perception visuelle :	17
4. La perception des couleurs :	18
4.1. La couleur :	18
5. Les systèmes de couleurs :	19
5.1. Le système colorimétrique RVB (ou RGB en anglais) :	21
5.2. Le système CIE XYZ :	23
5.4 Le système TSL /TSV (Teinte Saturation Luminance / Teinte Saturation Valeur) :	25
5.5. L'espace colorimétrique absolu CIE L*a*b 1976 :	25
6. Conclusion :	26
Chapitre 02 :	27
Méthodes et modèles de perception visuelle	27
1. Introduction :	28
2. L'attention et la perception visuelle :	28
2.1. L'attention visuelle :	28
2.2. La perception visuelle :	28
2.3. La perception et l'attention visuelle :	29
3. Les méthodes de perception visuelle :	29
3.1. Méthode ascendante (Bottom-up) :	29
3.2. Méthode descendante (Top-down) :	30
3.3. Bilan (Bottom-up, Top-down ou les deux) :	31
4. Théorie de recherche visuelle :	32
4.1. Théorie d'intégration des caractéristiques(FIT) :	32
4.1.1. Feature Search ou recherche disjonctif (recherche de caractéristiques) :	32
4.1.2. Conjonction Search (recherche de conjonction) :	33

4.2. Guided Search(GS) :	34
5. Carte de saillance :	35
5.1. La saillance visuelle :	35
5.2. La carte de saillance :	35
5.3. Les étapes de la carte de saillance :	36
6. Les modèles de la perception visuelle :	36
6.1. Le modèle de Koch et Ullman :	36
6.2. Modèle de Nervez Imamoglu & al :	37
6.3. Modèle de Steven Le Moan & al:	38
7. Conclusion :	39
Chapitre 03 :	40
La mise en œuvre du modèle d'Itti	40
1. Introduction :	41
2. Le modèle d'itti :	41
3. Les étapes du modèle :	42
3.1. Filtrage linéaire :	42
3.2. La pyramide gaussienne :	44
3.2.1. Définition	44
3.3. Différence centre-voisins :	45
3.4. La normalisation :	47
3.5. Combinaison des cartes :	48
3.6. Combinaison linéaire des cartes de visibilité (carte de saillance) :	48
4. Conclusion :	49
Chapitre 04	50
Utilisation du modèle d'Itti pour l'amélioration des images	50
(Implémentation et résultats)	50
1. Introduction :	51
2. Langages de programmation :	51
• C++ :	51
3. Bibliothèques utilisées :	52
• OpenCV :	52
4. Outils de developement :	52
• Visual Studio 2015 :	52
5. Modèle d'Itti pour l'amélioration des images :	53
5.1. Schéma globale :	53
5.2. Amélioration du contraste :	53

5.3. Les procédures et fonctions utilisées :	54
a. La 1ère étape : l'extraction des caractéristiques :	54
b. La 2ème étape : La différence centre voisin :	55
c. 3ème étape : La normalisation :	56
6. Résultats et Représentation des interfaces :	56

Liste des figures :

Chapitre 01: la perception visuelle et les système des couleurs

Figure (1) : l'œil humaine.....	14
Figure (2) : schéma de la rétine.	15
Figure (3) : schéma du « cône et bâtonnet »	16
Figure (4) : Construction de perception visuelle.....	18
Figure (5) : Spectre des ondes électromagnétiques et différents domaines d'application (extrait du cours de Physique Générale PhyG II-7).....	19
Figure (6) : Sensibilités (non normalisées) $L(\lambda)$, $M(\lambda)$ et $S(\lambda)$ des cônes L, M et S.....	20
Figure (7) : couleurs additives. (Les couleurs primaires rvb).....	21
Figure (8) : couleurs soustractives (couleurs primaire cmj « film, imprimante »).....	22
Figure (9) : Système RVB.	23
Figure [10]: Diagramme de chromaticité CIE (x, y). La coordonnée trichromatique x est en abscisse et y en ordonnée.....	24
Figure (11) : CIE XYZ (fonction d'observation d'un standard).....	24
Figure (12) : Schéma représente les modeles HSL/TSL(Hue/Teinte Saturation Luminance) et Le HSV ou TSV(Hue/Teinte Saturation Value/Valeur).....	25
Figure (13) : Espace $l^* a^* b^*$	26

Chapitre 02: Méthodes et modèles de perception visuelle.

Figure (1): "Buttom-up".....	30
Figure (2) : « Top-Down ».....	31
Figure(3) : feature search	33
Figure(4): Conjunition search.	34
Figure(5) : Guided Search.....	35

Figure(6): Le modèle de Koch et Ullman.	37
Figure(7): L'architecture globale du modèle de Nervez Imamoglu & al 2013. (Extrait de [Imamoglu et al., 2013]).....	38
Figure(8): Architecture globale du modèle Steven la Moan & Al.....	39
Chapitre 03 : la Mise en œuvre du modèle d'Itti	
Figure (1) : architecture globale de modèle d'itti.....	43
Figure(2): Exemple d'une pyramide.....	45
Figure(3) : normalisation des cartes	49
Figure(4): carte de saillance finale.....	50
Chapitre04 : Utilisation de la perception visuelle pour l'amélioration des images (implémentation et résultats)	
Figure (1): schéma globale du notre système	54
Figure (2) : interface principale.....	57
Figure (3) : Image d'entrée en rvb.....	58
Figure (4) : Carte d'intensité.....	59
Figure(5) : cartes des couleurs.....	60
Figure(6) : les cartes d'orientation	60
Figure(7) : Pyramide d'intensité	61
Figure(8) : carte de caractéristique d'intensité.....	62
Figure(9) : carte de caractéristique de couleur r_v	62
Figure(10) : carte de caractéristique de couleur b_j	63
Figure(11) : carte de visibilité d'intensité	63
Figure (12) : carte de visibilité de couleur.....	64
Figure(13) : carte de visibilité d'orientation.....	65
Figure(14) : carte de saillance	65
Figure(15) : accentuation du contraste et la luminosité sur l'image	66
Figure(16) : résultat finale.....	67

Introduction générale

La perception visuelle, en tant qu'outil de gestion de l'information ; nous permet de la modélisation des parties attirantes selon la vision humaine. La détection de saillance de l'image est la clé de l'extraction des informations d'image. L'extraction des régions de saillance d'image 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 d'image importants fournissent l'information la plus complète sur une image entière. Par conséquent, l'extraction des régions saillantes des images facilite efficacement de nombreuses applications d'image telles que la récupération d'image, la compression d'image adaptative, la reconnaissance d'objet, le redimensionnement d'image, « l'amélioration des images ».

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.

Notre objectif est d'exploiter les limites de la perception visuelle chez les humains pour trouver un compromis entre la qualité visuelle et le coût de calcul, ce qui permet de trouver des représentations optimales des scènes à rendre.

Dans ce mémoire, nous avons proposé d'étudier et d'implémenter un modèle de perception « Modèle d'Itti » 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.

Notre mémoire se compose de 04 chapitres :

Dans le chapitre 01 : nous allons présenter quelques propriétés de l'œil et de la perception visuelle, aussi que des notions sur la couleur et les systèmes de codifications existants, et la façon avec laquelle se fait la perception des couleurs.

Introduction générale

Dans le chapitre 02 : nous allons exposer les différentes méthodes et les modèles de perception visuelle, aussi qu'une étude théorique sur les techniques de la perception visuelle tel que le Bottom-up et le Top-down, nous parlera aussi sur l'attention visuelle et la carte de saillance qui sont principalement les éléments qui vont nous guider de voir et de focaliser sur une partie pertinente dans l'image.

Dans le chapitre 03 : nous allons détailler un modèle de perception visuelle « le modèle d'Itti », nous allons définir les objectifs de notre travail et les différentes étapes que nous allons entreprendre pour aboutir à la réalisation de notre application.

Le chapitre 04 concerne la partie implémentation, dans laquelle nous allons présenter le schéma global de notre projet, les différentes structures de données et algorithmes utilisés et enfin les résultats auxquels nous avons aboutis.

Chapitre 01 :

La perception visuelle et systèmes de couleurs

1. Introduction :

La perception visuelle est composée de deux composants majeurs : l'œil et le cerveau.

Nous allons commencer d'abord par introduire l'œil, ces composants et leur fonctionnement ; ensuite nous allons essayer de comprendre comment le cerveau traite-t-il ces informations, afin de faire une interprétation de l'image reçue.

Nous allons également parler de la perception des couleurs et les approches de la représentation des couleurs, qui sont un élément clé dans la problématique abordée par ce projet de fin d'études.

2. L'œil et son fonctionnement :

2.1. Définition :

L'œil est un organe extrêmement complexe, fruit d'une très longue évolution. C'est lui qui nous permet de voir le monde en couleur et de percevoir des milliers de couleurs. [1]

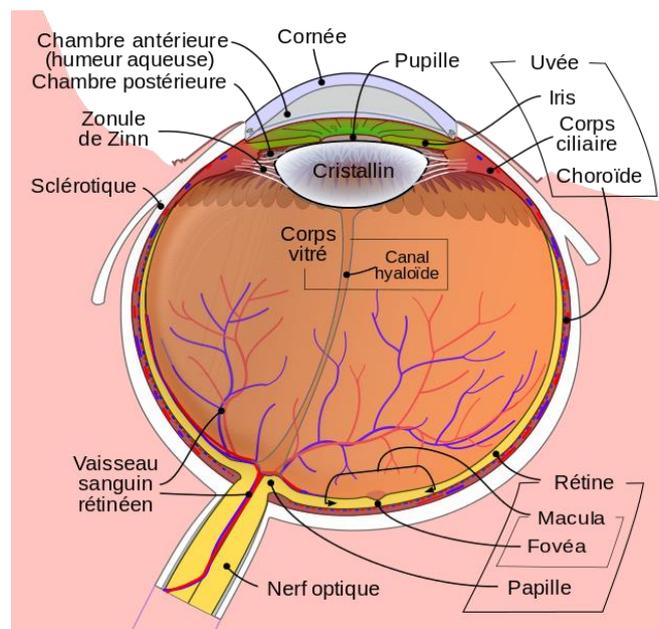


Figure (1) : l'œil humain. [2]

2.2. Composants de l'œil :

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

- a) **La sclérotique**, blanche et résistante, constitue le « blanc de l'œil ». Vers l'avant, cette enveloppe devient transparente et forme la cornée. [2]
- b) **La choroïde**, 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. [2]
- c) **La rétine** :

C'est l'enveloppe la plus interne de l'œil : c'est une membrane très mince et transparente dont le prolongement forme le **nerf optique**, c'est la composante de la perception visuelle. [2]

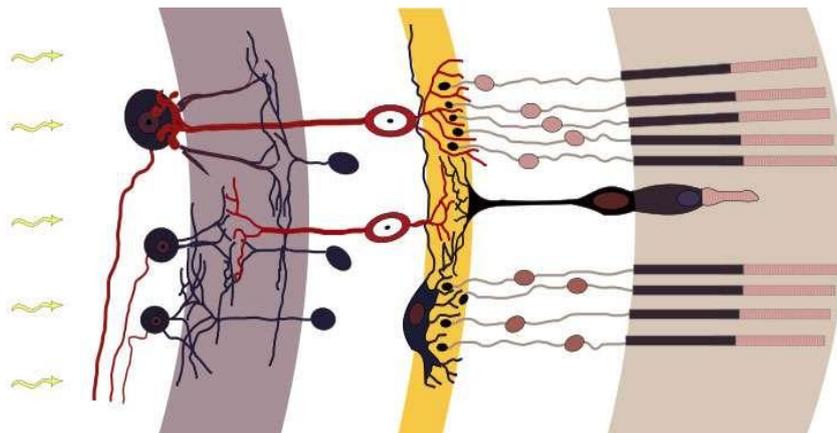


Figure (2) : schéma de la rétine. [4]

Au niveau de la rétine il existe deux types de photorécepteurs se sont : Les cônes et les bâtonnets ; ce sont des neurones très courts sensibles à la lumière et caractérisées par leur forme et la nature des pigments.

Il existe 7 millions de cônes, et de 130 millions du bâtonnet.

❖ Les cônes :

Les cônes sont essentiellement situés dans la zone centrale de la rétine « la fovéa », ils permettent une vision précise et une discrimination des couleurs mais exige une intensité lumineuse assez forte (vision du jour). [3]

Il y a 3 types de cônes sensibles à la lumière qui sont : le S M L

- S : Short : ces cônes sont sensibles aux couleurs bleus.
- M : Medium : ces cônes sont sensibles aux couleurs verts.
- L : Long : ces cônes sont sensibles aux couleurs rouges. [4]

❖ **Les bâtonnets :**

Ils fonctionnent avec très peu de lumière, et sont utilisés pour la vision nocturne (dans la nuit).

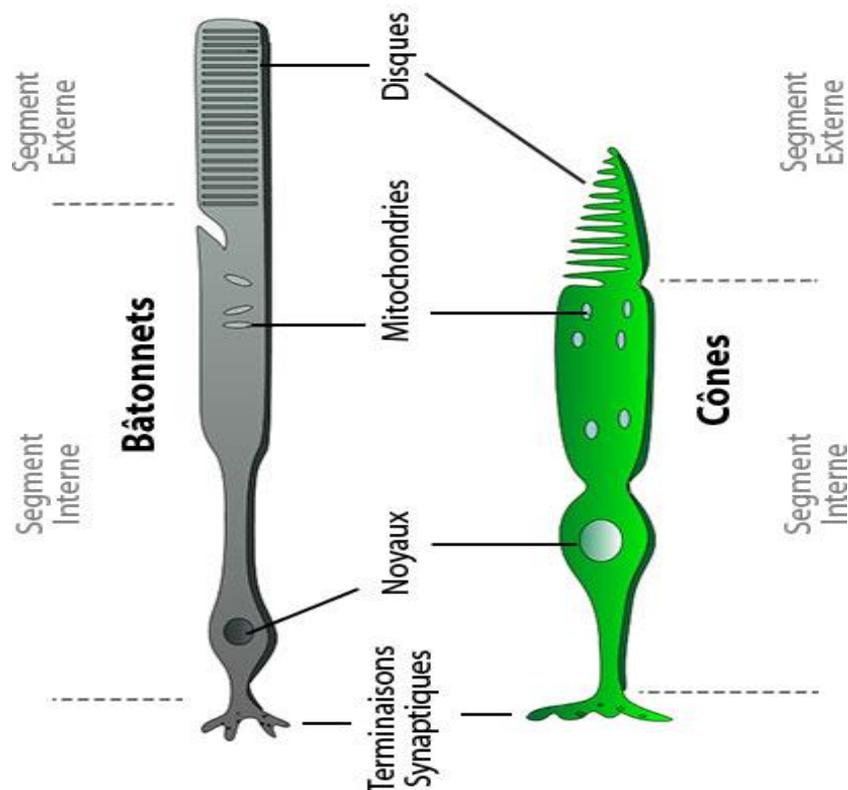


Figure (3) : schéma du « cône et bâtonnet » [7]

L'intérieur des globes oculaires est rempli de différents **milieux transparents** qui, en jouant le rôle d'une lentille, permettent la formation d'une image au fond de l'œil ; il s'agit de l'humeur aqueuse, du cristallin et de l'humeur vitrée.

Entre l'iris et la cornée se trouve un espace rempli d'**humeur aqueuse**, un fluide semblable à l'eau. [2]

Derrière l'**iris** se trouve le **cristallin**, une lentille transparente biconvexe qui permet de faire la « mise au point » pour des objets rapprochés. Enfin, l'espace situé entre le cristallin et la rétine contient l'**humeur vitrée**. [2]

3. 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]

3.1. Construction de la perception visuelle :

- ✚ Les réactions chimiques ayant lieu dans les cônes et les bâtonnets entraînent la création de messages nerveux qui se propage dans le nerf optique ; ce sont des signaux électriques dont la fréquence correspond à une information codée.
- ✚ Ces messages nerveux empruntent les voies visuelles pour atteindre un cortex visuel occipital à la partie arrière du cerveau qui traite les informations visuelles.
- ✚ Les messages issus des parties temporales de chaque rétine se propagent sans changer l'hémisphère cérébral.
- ✚ Les messages nerveux issus de la partie nasale de chaque rétine se croisent au niveau des « chiasma optique » et rejoignent le cortex visuel occipitale de l'hémisphère opposé.
- ✚ Au cours de leur trajet, les messages passent d'une fibre nerveuse à une autre au niveau de zone de jonction appelé « synapse », là les signaux électriques qui arrivent de la rétine engendrent la libération de substance chimique appelé « neurone transmetteur » qui traverse l'espace entre les deux fibres nerveuses et qui se fixe sur des récepteurs de forme complémentaire, cette fixation entraîne la naissance de nouveaux signaux sur la fibre reliée au cortex.
- ✚ Le cortex visuel occipital est constitué de différentes aires spécialisées dans le traitement des messages nerveux selon la fréquence des signaux électriques générées par les photorécepteurs, chaque aire interprète une caractéristique de l'image observée :
Forme, Couleurs, Mouvement.
- ✚ Ces différentes aires communiquent entre elles et le cerveau réalise alors une interprétation globale.

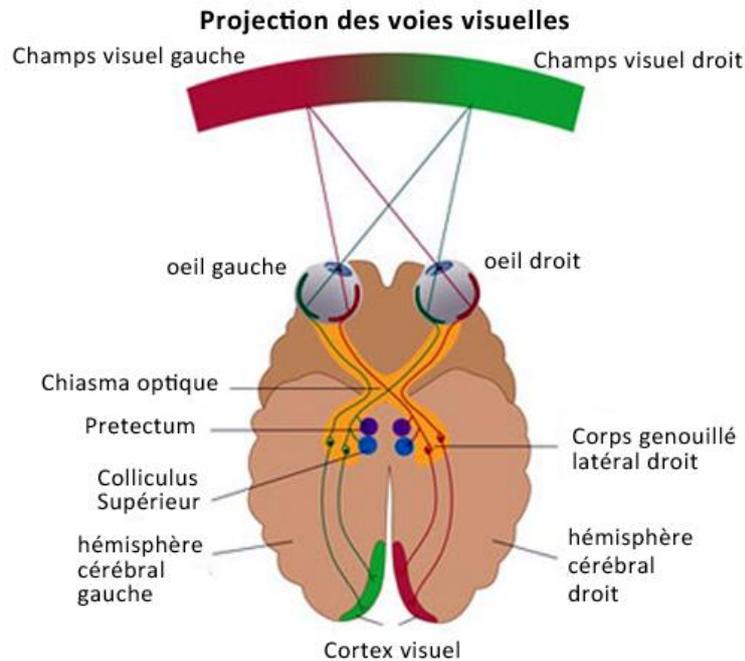


Figure (4) : Construction de perception visuelle. [8]

4. La perception des couleurs :

4.1. La 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.[9]

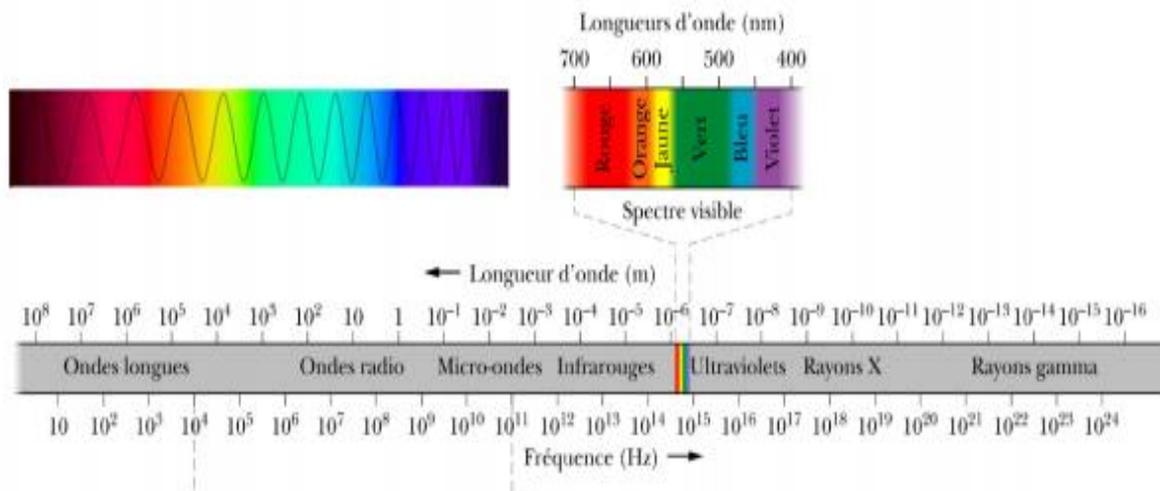


Figure (5) : Spectre des ondes électromagnétiques et différents domaines d'application (extrait du cours de Physique Générale PhyG II-7). [10]

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. [10]

5. 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.

Comme nous venons de le voir, l'œil comporte trois types de cônes dont les courbes de sensibilité sont données par la figure suivante :

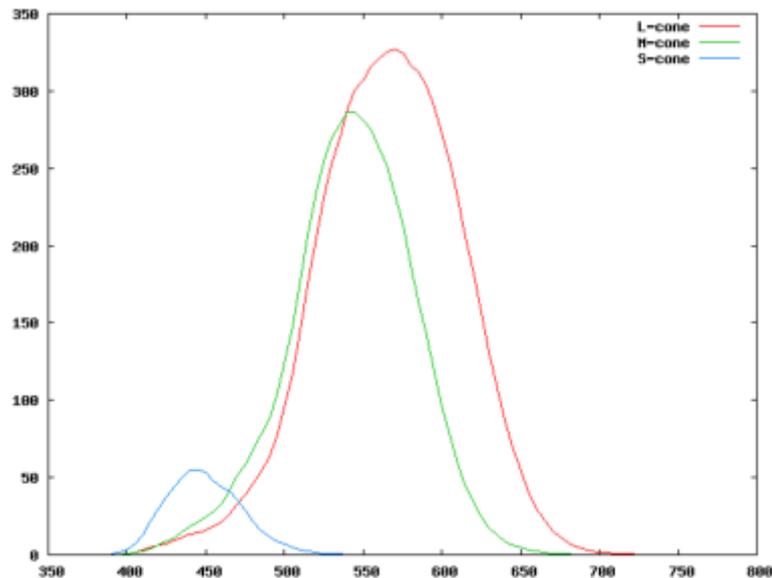


Figure (6) : Sensibilités (non normalisées) $L(\lambda)$, $M(\lambda)$ et $S(\lambda)$ des cônes L, M et S

Cette figure trace les sensibilités des trois types de cône pour une même quantité d'énergie reçue en fonction de la longueur d'onde. On voit que les cônes bleus sont moins sensibles que les cônes verts et rouges.

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. [11]

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** : L'« 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.



5.1. Le système colorimétrique RVB (ou RGB en anglais) :

C'est un système de codage informatique des couleurs les 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.

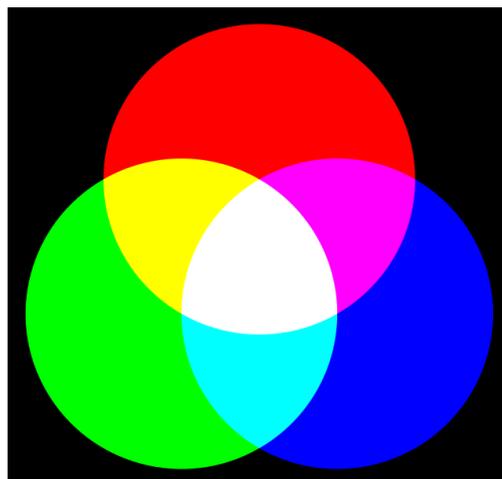


Figure (7) : couleurs additives. (Les couleurs primaires rvb) [11]

- **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.

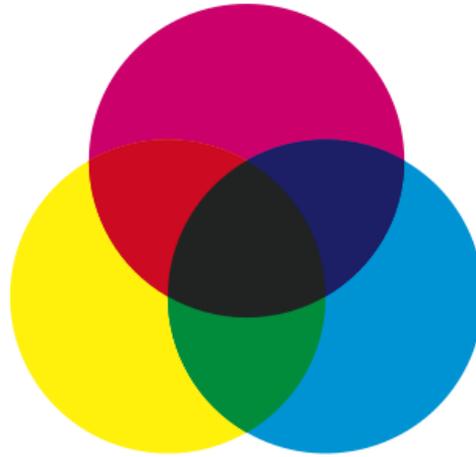


Figure (8) : couleurs soustractives (couleurs primaire cmj « film, imprimante ») [11]

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. [11]

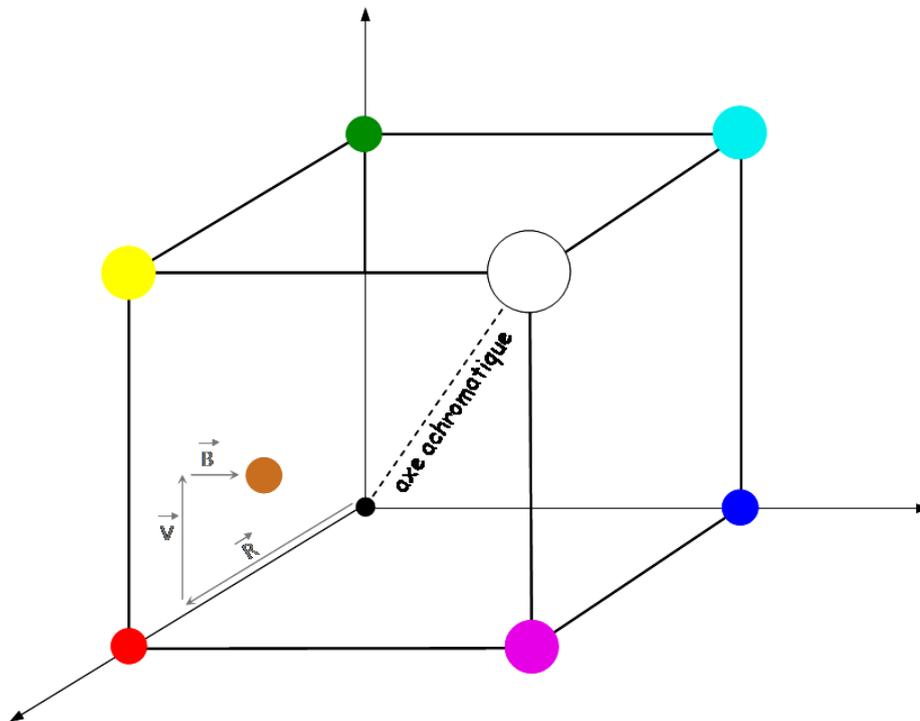


Figure (9) : Système RVB. [11]

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.

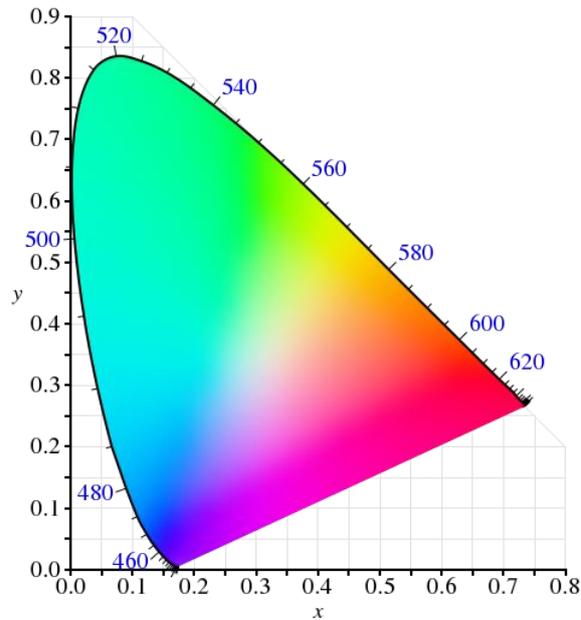
Les points de cette droite représentent des nuances de gris allant du noir au blanc. Elle a pour équation $R_c = G_c = B_c$. [11].

5.2. 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[10]: 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 définit 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.

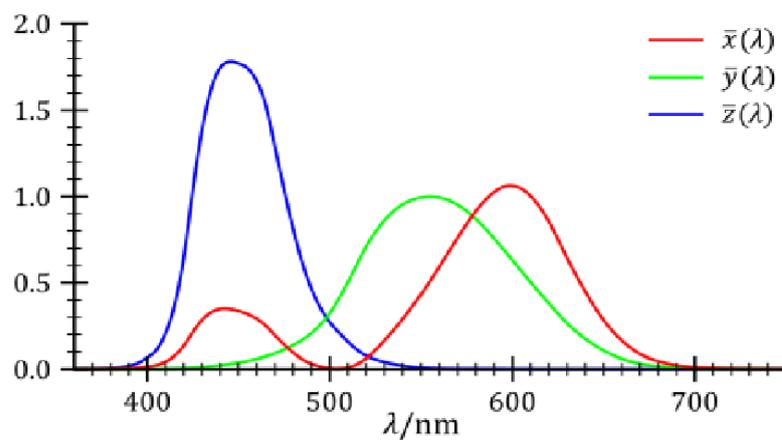


Figure (11) : CIE XYZ (fonction d'observation d'un standard). [11]

5.4 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 %.

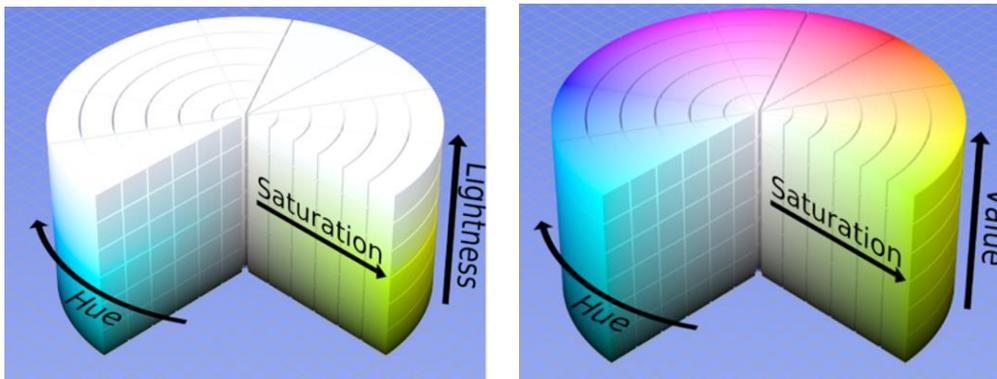


Figure (12) : Schéma représente les modèles HSL/TSL (Hue/Teinte Saturation Luminance) et Le HSV ou TSV (Hue/Teinte Saturation Valeur/Valeur). [11]

5.5. 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 :

L'utilisation de la perception visuelle pour l'amélioration des images

- Représentation en coordonnées rectangulaires $L^*a^*b^*$.
- Représentation en coordonnées cylindriques $L^*C^*h^\circ$.

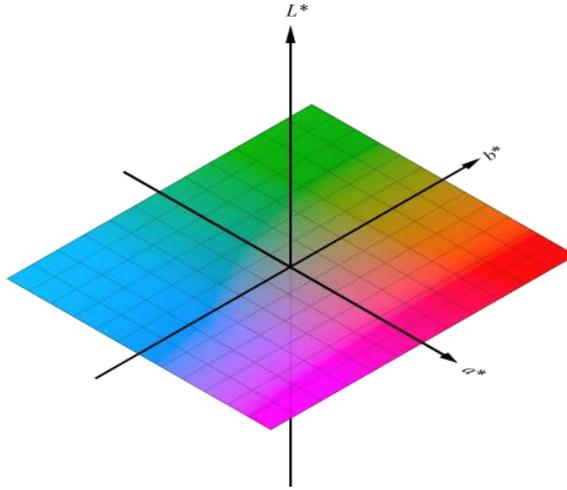


Figure (13): Espace $l^ a^* b^*$. [11]*

6. Conclusion :

Dans ce chapitre nous avons défini le fonctionnement de l'œil qui est primordial pour la suite de notre étude, ensuite nous avons étudié la relation entre l'œil et la perception visuelle, et enfin la perception des couleurs et la représentation des couleurs. Cette étude nous a permis de comprendre les mécanismes essentiels liés à la perception visuelle.

Dans le chapitre suivant nous allons parler de mécanismes et modèles de perception visuelle pour la représentation des scènes.

Chapitre 02 :

Méthodes et modèles de perception visuelle

1. Introduction :

L'étude du système visuel humain a permis de comprendre quelques aspects de son fonctionnement. Plusieurs processus complexes interagissent entre eux et agissent sur l'information visuelle depuis son acquisition par l'œil jusqu'à son interprétation dans le cortex visuel. Les résultats des recherches dans ce domaine ont donné naissance à une diversité de modélisations et des simulations de ce système.

Modéliser la vision a toujours été un défi en informatique. Il apparaît qu'étudier le cerveau et la vision humaine peut faire progresser en vision par ordinateur.

Dans ce chapitre nous allons définir quelques modèles et méthodes de perception visuelle, puis quelques termes importants dans la perception visuelle tel que la saillance et la carte de saillance.

2. L'attention et la perception visuelle :

2.1. L'attention visuelle :

L'attention est l'action de concentrer l'esprit sur un ou plusieurs éléments particuliers. Cela renvoie au fait que nous ne traitons pas avec le même degré de profondeur l'ensemble des informations en provenance de notre environnement. Ainsi, le déplacement du regard ne se fait pas de manière aléatoire sur l'image mais suit les traits saillants, privilégiant certains aspects de la scène. Le regard reste en général environ un quart de seconde sur un point, la scrutation à proprement parler ne durant que 50 ms.

Le rôle principal de ce mécanisme est de réduire sensiblement la quantité d'informations visuelles qui sera traitée par des tâches complexes, telle que la reconnaissance d'objets, entraînant ainsi l'accélération de l'ensemble du processus de la vision. [15]

2.2. La perception visuelle :

C'est toute sensation intérieure de connaissance apparente, résultant des stimuli ou d'une impression lumineuse captée par les yeux (la vue).

En général, cet acte optique et physique fonctionne plus ou moins pareil chez toutes les personnes, étant donné que les différences physiologiques des organes visuelles n'affectent que le résultat de la perception.

2.3. La perception et l'attention 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. La réception des stimuli visuels est effectuée par l'œil (la rétine en particulier « chapitre1 »).

La quantité d'informations provenant de ces stimuli visuels étant trop importante pour être traitée dans sa totalité, un mécanisme de sélection des informations est nécessaire : c'est le rôle de l'**attention visuelle**.

3. Les méthodes de perception visuelle :

3.1. Méthode ascendante (Bottom-up) :

Les processus Bottom-up, appelé encore ascendant sont des mécanismes exogènes guidée uniquement par les stimuli présents du champ visuel sans aucune volonté de la part de l'observateur qui définit la notion de saillance contient une idée centrale, Un objet est saillant s'il est facilement remarqué. A partir de cette idée intuitive, plusieurs définitions peuvent être proposées, selon le critère qui fait que l'objet va être remarqué. la saillance s'appuient uniquement sur le contenu de l'image et les régions saillante d'une image sont donc les même quel que soit l'observateur.

Carte de saillance : sont généralement obtenues à partir de caractéristiques visuelles simples, applicables sur l'ensemble de l'image à traiter. Permettent de regrouper au sein d'une image différentes information photométriques et colorimétriques (contrastes de luminance, contrastes colorés, bords, angles, etc.)

Avantage

→ Le processus Bottom-up est plus rapide et qu'il précède des influences top-down plus longue à mettre en œuvre et qui va durer dans le temps

Inconvénients

- Elle ignore les informations sémantiques, qui ont pourtant une influence sur ce qui est vu et ce qui ne l'est pas.
- Elle ne permet pas d'estimer la saillance des objets, mais seulement de régions de l'espace (concrètement, des pixels de l'image). Dans une approche par analyse d'images, le passage du niveau des pixels au niveau des objets implique, concrètement, une phase de segmentation des objets dans l'image. [14]

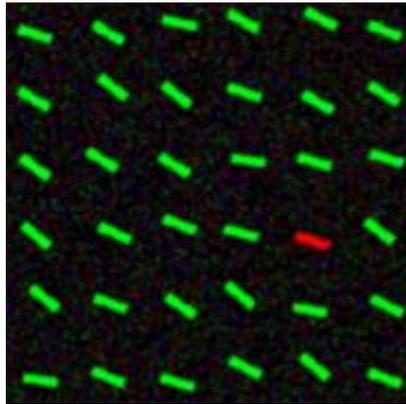


Figure 1 : "Bottom-up".

La figure 1 montre qu'un seul objet rouge dans un champ vert sera saillant et attire notre attention d'une manière.

3.2. Méthode descendante (Top-down) :

Les méthodes descendantes sont des mécanismes endogènes faisant intervenir la volonté du sujet. Ce sont des processus contrôlés. Dans certains cas c'est le cerveau lui-même qui envoie directement l'information vers les systèmes sensoriels. En effet les perceptions peuvent être influencées par ce que s'attend à voir ou a en mémoire une personne. Ainsi l'approche top-down est une approche principalement dirigée par les concepts. C'est plutôt la connaissance a priori qui va guider le processus de vision

Les indices de cette méthode sont déterminés par les phénomènes cognitifs (reconnaissance, prédiction, ...). [13]

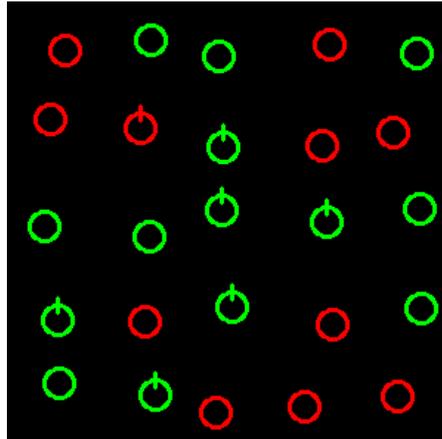


Figure2 : « Top-Down »

Comme il est montré dans la figure 2 il est difficile de connaître le cercle rouge avec un trait rapidement

3.3. Bilan (Bottom-up, Top-down ou les deux):

Le système attentionnel humain est influencé par deux sources d'information : endogène (Bottom-up) et exogène (Top-down). Un modèle artificiel de perception visuelle devrait donc prendre en compte ces deux composantes.

Cependant, la partie top-down est difficile à appréhender car dépendant, par sa définition, du contexte. C'est certainement pour cette raison que la grande majorité des modèles de la fin des années 1990 et du début des années 2000 ne traitait que de la composante Bottom-up. On notera cependant quelques exceptions avec des modèles Bottom-up et top-down très proches de la théorie, mais aussi avec des approches plus computationnelles.

Dans ce dernier cas, l'implémentation du mécanisme top-down consiste, soit en la possibilité de modifier globalement ou localement le poids de certaines cartes, soit au calcul d'une carte top-down à partir d'un système de reconnaissance d'objets.

De manière assez logique, aucun modèle n'est purement top-down, puisque cette influence est considérée comme une modulation contextuelle de l'attention bottom-up (guidée par les données).

L'ajout d'une composante top-down permet donc de créer un système de perception supervisé, mieux à même de s'adapter à un contexte d'utilisation spécifique.

Ainsi, une grande partie des modèles récents propose d'intégrer les deux sources d'attention, la partie top-down pouvant être prise en compte en utilisant : « une modification globale ou locale des poids de chaque caractéristique (intensité, couleur, orientation, etc.) en fonction des propriétés de l'objet à reconnaître ».

Cela peut être réalisé soit en fournissant directement une ou des cartes de caractéristiques correspondant à la cible, ou par un processus d'apprentissage permettant de calculer les poids séparant aux mieux la cible du fond. [15]

Nous concentrons dans notre projet sur le modèle bottom-up par ce qu'il détecte les informations facilement du contraire de méthode Top-down qu'elle besoin trop de temps de calcul.

4. Théorie de recherche visuelle :

A chaque instant nous cherchons des objets (les mots de ce texte, un stylo sur la table, une tasse...) soit de manière explicite, en bougeant les yeux, soit de manière implicite, en déplaçant la cible de notre attention. La manière générale, dans une tâche de recherche visuelle, les sujets ont pour consigne de rechercher un objet cible ou (objet saillant) prédéterminé parmi un ensemble d'objets non cibles, dits aussi distracteurs.

4.1. Théorie d'intégration des caractéristiques(FIT):

4.1.1. Feature Search ou recherche disjonctif (recherche de caractéristiques) :

La primauté de la perception des composantes transparaît dans la théorie de l'intégration des traits de Treisman : la vision précoce consisterait à extraire de l'image des traits, alors représentés sur des plans distincts (*feature maps*), en concordance avec la spécialisation des aires corticales impliquées en vision.

C'est pourquoi la recherche visuelle d'une cible à partir du critère d'un trait unique sur lequel repose sa différence avec des distracteurs (*e.g.*, la cible Q parmi des O) est facilement réalisée, grâce à un phénomène de 'pop-out':

La cible est perçue immédiatement, sans effort et automatiquement, avec une même facilité quelle que soit la quantité de distracteurs.

La détection immédiate de la cible est alors réalisée de manière pré-attentionnelle, grâce à un mécanisme de traitement en parallèle. De même, dans une scène visuelle, une rupture au niveau d'une dimension unique telle que la couleur ou la luminance permettrait de délimiter automatiquement une forme parce que l'extraction d'aucune relation entre traits n'est nécessaire. [20]

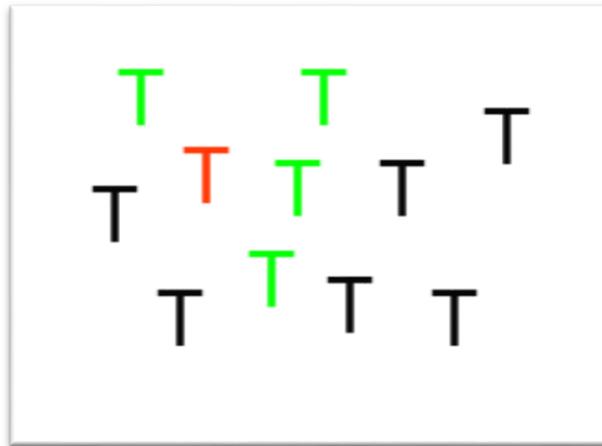


Fig3 : feature search « il y a un « T » on couleur rouge dans la collection »

4.1.2. Conjonction Search (recherche de conjonction) :

Par contre, si la distinction de la cible nécessite la prise en compte d'une conjonction de traits relevant de dimensions différentes (*e.g.*, *le cible est définie par la forme et la couleur, ou Y a-t-il un T rouge dans cette affichage .figure4*), sa détection devient plus difficile et prend d'autant plus de temps que le nombre de distracteurs augmente. Selon la théorie de l'intégration des traits, il s'impose alors de prendre en compte des relations entre traits pour les intégrer (*binding*), ce qui impliquerait des opérations coûteuses en temps et en énergie. [20]

Ces phénomènes de conjonction illusoire constituent un argument important pour la conception analytique de la perception (Printzmetal, 1995). Ils se produisent surtout à partir de la similarité de couleur. Une combinaison correcte des différents traits serait donc une étape de traitement attentionnelle et secondaire au codage séparé des traits.

Comme le champ visuel ne serait pas représenté sous forme d'objets dans le traitement pré-attentionnel, mais comme une collection de traits, l'attention ne serait pas dirigée par des objets (*object-based attention*), mais sélectionnerait plutôt des régions de l'espace (*space-based attention*) pour permettre seulement ensuite la construction mentale des objets. Avant cet assemblage, le monde visuel se limiterait à une 'soupe' d'attributs sans organisation spatiale (Rousselet, 2003).

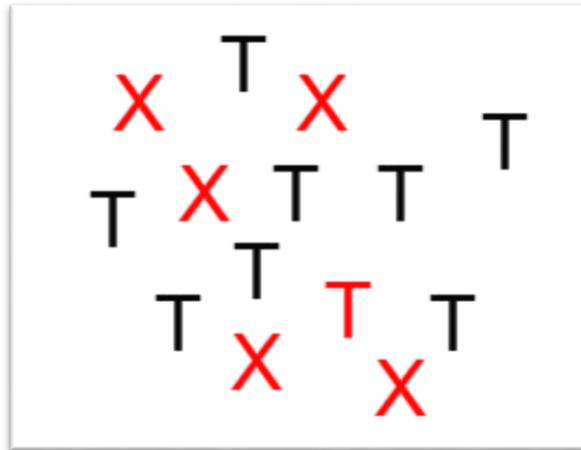


Fig4 : Conjunction search.

4.2. Guided Search(GS) :

Dans le modèle de la recherche guidée (*guided search theory*, Cave & Wolfe, 1990), la stricte distinction entre recherches en parallèle ou séquentielle est abandonnée : pour rechercher une cible définie par une conjonction de traits, des processus visuels de bas niveau extrairaient les traits élémentaires en parallèle pour différentes dimensions, puis un mécanisme de guidance qualifié de 'pré-attentionnel' utiliserait cette information pour diriger l'attention sur les items possédant l'un et/ou l'autre trait pour décider s'il s'agit ou non de la cible.[23]

Toutefois, selon ce modèle, les configurations ne semblent pas constituer des primitives guidant par elles-mêmes la recherche. Le système visuel ne découperait le monde qu'en 'proto-objets', sortes de listes de propriétés ou objets potentiels non encore définis comme de véritables configurations :

les mécanismes visuels précoces ne feraient qu'isoler des listes de propriétés témoignant de leur présence, mais pas de leur organisation et de leur forme (Wolfe & Bennett, 1997) par exemple la figure 6 montre :que le processus de recherche séparé pour les Xs et les objet blancs

les Os (caractéristiques cibles), et il y a une double activation qui dessine l'attention pour le cible.

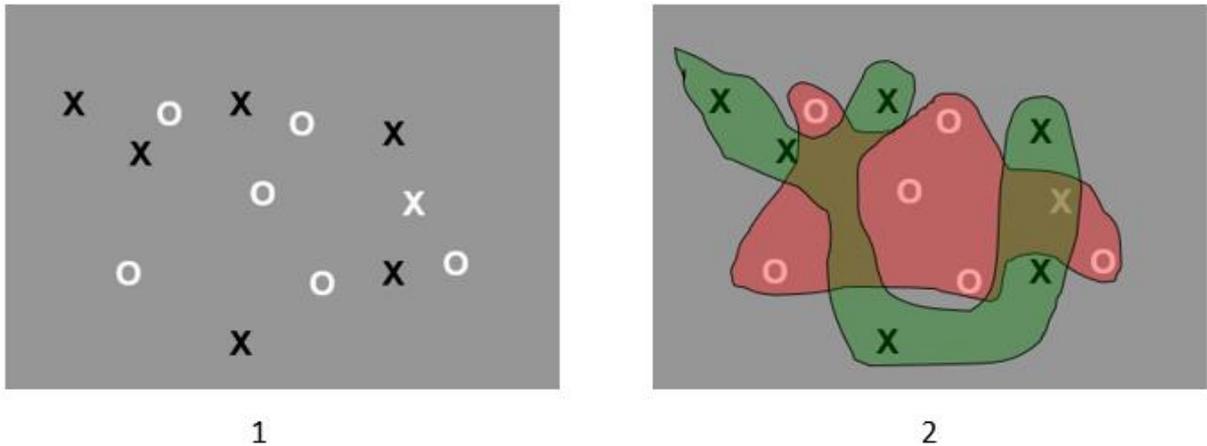


Fig5 : Guided Search

5. Carte de saillance :

5.1. La saillance visuelle :

La saillance visuelle est la qualité de perception subjective distincte qui rend un objet remarquable par rapport à ses voisins et qui attire immédiatement notre attention.

Notre attention est attirée par des stimuli visuels saillants. Il est important pour les systèmes biologiques complexes de détecter rapidement les objets et les régions dans un champ de vision donné. Cependant, d'identifier simultanément toutes les cibles intéressantes dans un champ visuel provoque une complexité de calcul prohibitive, ce qui est en fait une tâche ardue même pour les cerveaux biologiques les plus sophistiquées, et encore moins pour n'importe quel ordinateur existant. Une solution adoptée par les primates et d'autres animaux, est de limiter le processus complexe de reconnaissance d'objets à une petite zone, ou peu d'objets à un moment donné. Les nombreux objets ou zones de la scène visuelle peuvent ensuite être traité l'une après l'autre. [16]

5.2. La carte de saillance :

Donc, la carte de saillance c'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 stimuli (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.

5.3. Les étapes de la carte de saillance :

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. [13]

6. Les modèles de la perception visuelle :

6.1. Le modèle de Koch et Ullman :

Ce modèle, proposé par Koch et Ullman 1985, l'attention visuelle est guidée par la combinaison des caractéristiques de bas niveau comme l'intensité de luminosité, la couleur et l'orientation. Ainsi, dans ce premier modèle de Koch and Ullman (Figure 6) :

- ✚ Une image d'entrée est décomposée en plusieurs cartes, une carte par caractéristique de bas niveau (Modèle Bottom up).
- ✚ Ensuite, dans ces cartes, les positions saillantes émergent en supposant que la saillance d'une position dépend de sa différence par rapport aux positions voisines.
- ✚ Finalement, les cartes de caractéristiques sont sommées pour créer la carte unique qui s'appelle la carte de saillance ("saliency map").

Cette carte de saillance, combinée avec le mécanisme WTA (“Winner-Take-All”), permet de prédire les positions que les sujets fixent.

Le maximum de la carte correspondra à la prédiction de la première fixation, puis le maximum suivant correspondra à la prédiction de la deuxième fixation, etc. Désormais, la notion de carte de saillance est largement utilisée dans des études concernant le modèle d’attention visuelle.

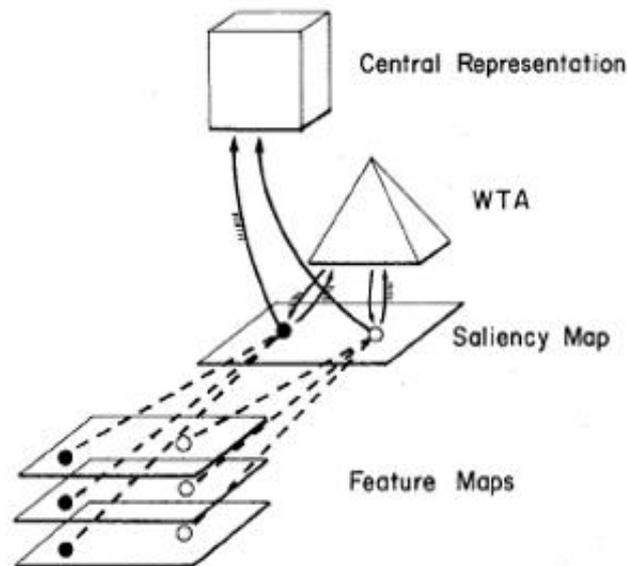


Figure6: *Le modèle de Koch et Ullman.*

Le modèle de Koch et Ullman joue un rôle important en présentant le modèle de base sur lequel s’appuient de nombreux modèles d’attention visuelle.

Les modèles développés à partir du modèle de Koch et Ullman apportent des améliorations au niveau de l’implémentation mais conservent l’architecture générale de celui-ci.[17]

6.2. Modèle de Nervez Imamoglu & al :

Le modèle est proposé par Nervez Imamoglu, Weisi Lin et Yuming Fang, publié en 2013. Il utilise les caractéristiques de bas niveau obtenues à partir de la transformée en ondelettes.

Les cartes de caractéristiques sont créées en augmentant les bandes passantes ou les composantes de fréquence. Comme il est montré dans l’architecture globale(fig6), le modèle intègre deux cartes de saillance : locale et globale. Les deux cartes sont obtenues par des caractéristiques de même niveau, basés sur les coefficients d’ondelette.

Les étapes pour obtenir la carte de saillance :

- Conversion RGB vers CIELab.
- Génération des cartes de caractéristiques.
- Calcul de la carte de saillance locale.
- Calcul de la carte de saillance globale.
- Fusionnement des cartes de saillance locale et globale.[17]

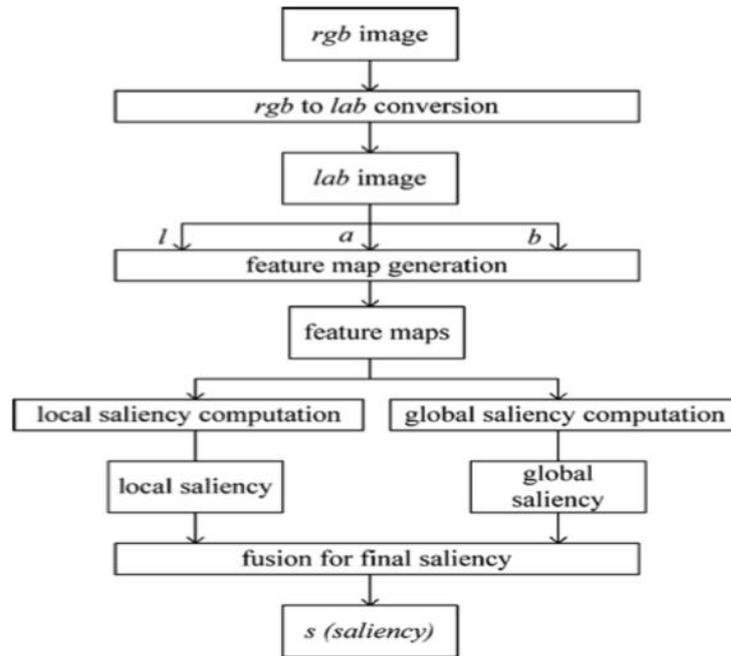


Fig7: L'architecture globale du modèle de Nervez Imamoglu & al 2013. (Extrait de [Imamoglu et al., 2013])

6.3. Modèle de Steven Le Moan & al:

Ce modèle de saillance spectral (SS) proposé par Steven Le Moan, Alamin Mansouri, Jon Y. Hardeberg et Yvon Voisin [Le Moan et al., 2013].

Ils introduisent un nouveau modèle d'extraction de caractéristiques utile pour la comparaison d'images, la visualisation et l'interprétation.

La notion de saillance spectrale est définie, comme une mesure dans laquelle un groupe de pixels se distingue dans une image et en termes de réflectance, plutôt qu'en terme d'attributs colorimétriques comme c'est le cas dans les études traditionnelles de saillance. Le modèle prend comme entrée une image multi-spectrales ou hyper-spectrales avec n'importe quelle dimension,

n'importe quelle gamme de longueurs d'onde, et il utilise une série d'extraction de caractéristiques pour délivrer en sortie une carte de saillance.

Le modèle SS est basé sur la comparaison de chaque pixel avec ces voisins en termes de différentes caractéristiques et en plusieurs échelles.

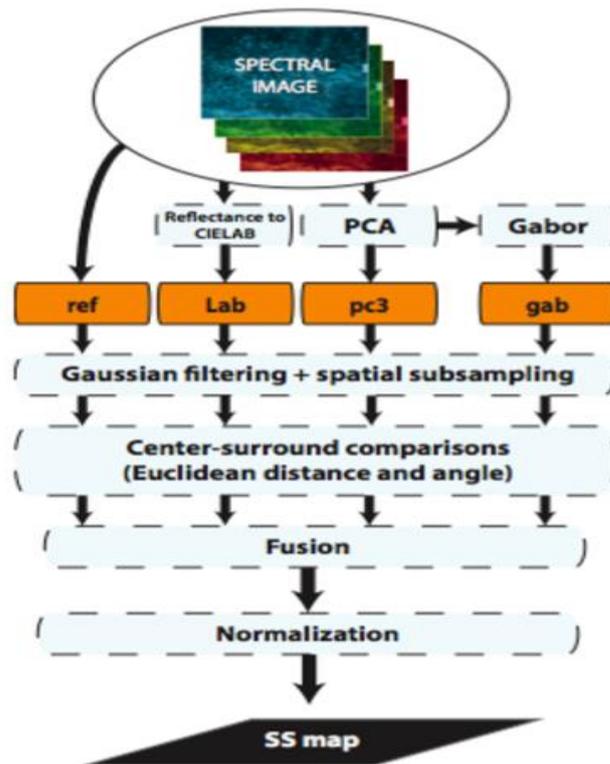


Figure8 : Architecture globale du modèle Steven la Moan & Al

7. Conclusion :

Dans ce chapitre, nous avons expliqué les différents modèles et méthodes de perception visuelle, ainsi que les concepts de saillance visuelle et la carte de saillance qui ce sont des éléments essentiels pour la perception visuelle.

Le chapitre suivant, concerne la partie de réalisation et mise en œuvre d'un modèle de perception visuelle pour le traitement d'image.

Chapitre 03 :

La mise en œuvre du modèle d'Itti

1. Introduction :

Nous avons vu dans le chapitre précédant les différents modèles de perception visuelle pour la modélisation d'une scène ou d'une image pour obtenir une carte de saillance qui sélectionne les partis attirants dans une scène quelconque,

Dans ce chapitre nous allons nous concentrer sur le modèle d'Itti ; qui est le modèle le plus cité et le plus connu.

2. Le modèle d'itti :

Le modèle d'Itti est proposé par Koch et Ullman en 1985 [Koch et Ullman, 1987], extrait au 1989. Il est inspiré par le comportement et l'architecture neuronale du système visuel des primates. Il devient une référence à des nombreux modèles par la suite, et sans doute c'est le modèle le plus cité et le plus connu.

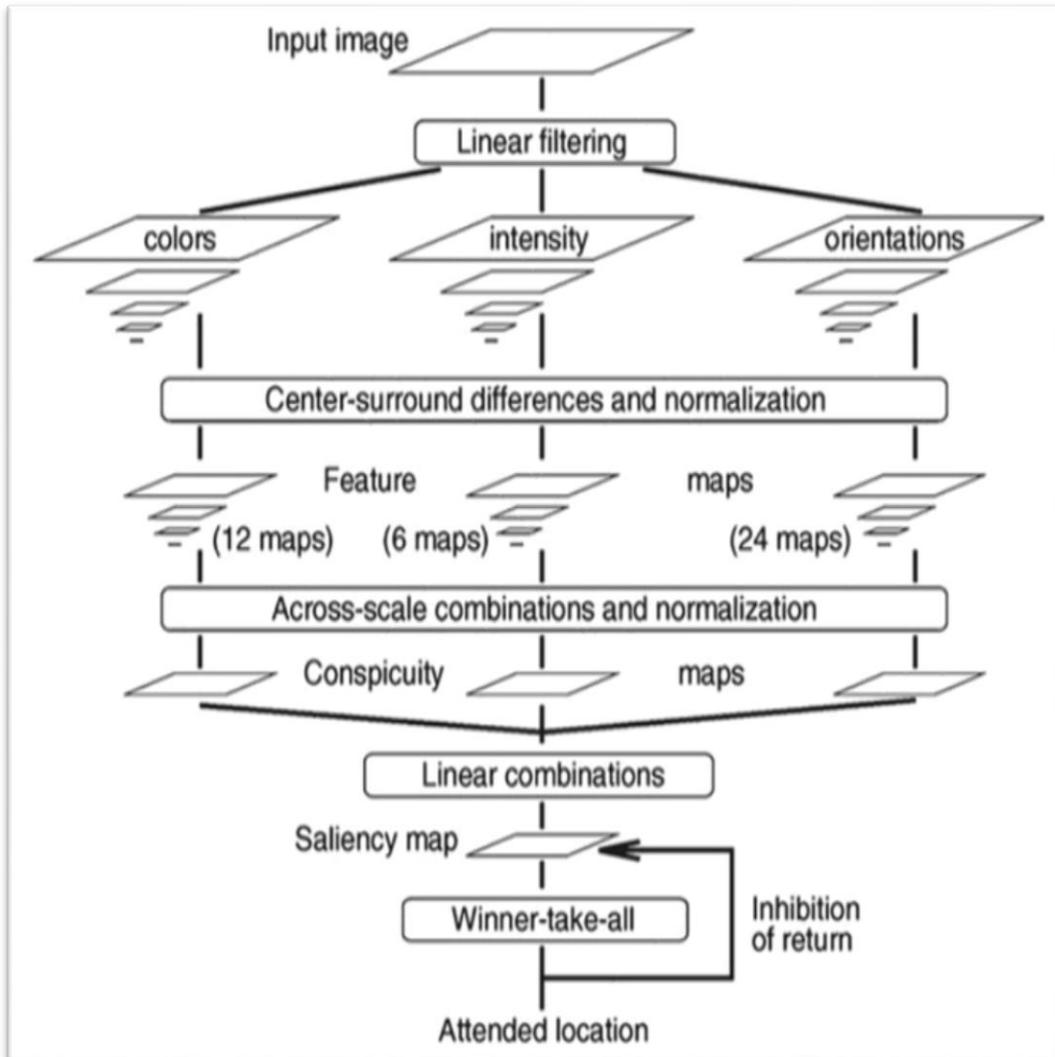
Ce modèle combine les caractéristiques multi-échelles de l'image vers une carte de saillance topographique unique. Il s'attache uniquement aux processus Bottom-up.

Comme le modèle original de Koch et Ullman, le modèle d'Itti décompose un stimulus visuel en **caractéristiques visuelles** de bas niveau : l'orientation, l'**intensité** et la **couleur**.

Les étapes du modèle d'Itti sont :

- ✚ Prétraitement visuel (l'extraction des caractéristiques)
- ✚ Différence centre-voisins.
- ✚ Normalisation.
- ✚ Cartes de visibilité.
- ✚ Carte de saillance

Le modèle d'Itti c'est un modèle qui permet d'extraire une carte de saillance d'une scène visuelle correspondante.



Figure(1) : architecture globale de modèle d'itti

3. Les étapes du modèle :

3.1. Filtrage linéaire :

Chez l'homme, l'information visuelle est traitée dans la rétine par des cellules achromatiques et chromatiques, sensibles respectivement à l'intensité et aux contrastes rouge / vert et bleu / jaune. Nous devons donc convertir les images RGB que notre système perceptuel reçoit en entrée dans un espace de couleur le plus adapté.

Le modèle d'Itti utilise l'espace RGB utilisé, Une fois la conversion des couleurs effectuée, nous disposons d'un canal achromatique I et de deux canaux chromatiques $I R/G$ et $I B/Y$. Ceux-

L'utilisation de la perception visuelle pour l'amélioration des images

ci vont servir de base au calcul des cartes de caractéristiques d'intensité, couleur, orientation via une approche multi-résolutions.

- **Carte d'intensité :**

Correspond à la valeur moyenne des trois canaux r, g, b représentés dans l'espace RGB

L' image d'intensité I se calcule de la manière suivante :

$$I = (r + g + b) / 3.$$

- **Carte de couleurs :**

Chaque couleur est calculé comme suit :

$$\text{Rouge } R = r - (g + b)/2,$$

$$\text{Vert } G = g - (r + b)/2$$

$$\text{Bleu } B = b - (g + r) / 2$$

$$\text{Jaune } Y = (g + r) / 2 - |r - g| / 2 - b.$$

- **Carte d'orientation :**

Est obtenu à partir de la carte d'intensité I en utilisant des filtres de Gabor $O(\sigma, \theta)$ (filtres passe-bande orientés ; ces filtres, dont le fonctionnement est proche des traitements visuels humains, ont l'avantage d'être paramétrables en fréquence, et en orientation) où $\sigma \in [0..8]$, qui représentent d'une manière plausible le champ récepteur des neurones sélectifs en orientation dans le cortex visuel primaire, elle permet de prédire les mouvements oculaires de l'œil durant les premiers secondes de vision.

Les orientations choisies sont : $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ, \}$.

$$h(x, y) = \exp\left(-\frac{x_0^2 + \gamma y_0^2}{2\sigma^2}\right) \cos\left(\frac{2\pi}{\lambda} x_0\right)$$

$$\mathbf{x}_0 = x \cos(\theta) + y \sin(\theta), \mathbf{y}_0 = -x \sin(\theta) + y \cos(\theta)$$

- Pour chaque carte on va appliquer la pyramide gaussienne, alors chaque carte de caractéristiques représente une pyramide

3.2. La pyramide gaussienne :

3.2.1. Définition :

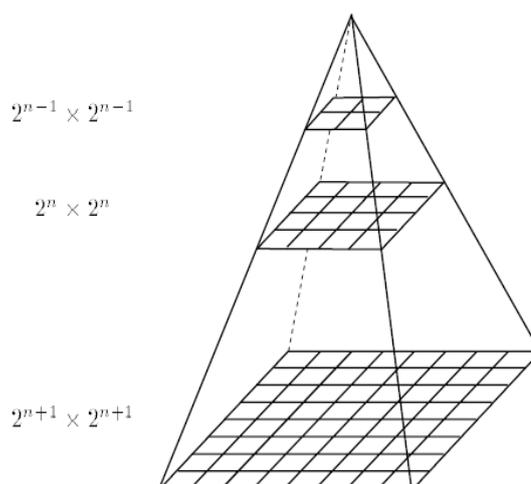
La pyramide gaussienne est une technique de traitement d'image qui décompose une image en plusieurs groupes successifs de pixels, en plusieurs étapes, dans le but de la rendre floue. Il est nommé d'après le mathématicien allemand *Johann Carl Friederich Gauss*.

Ce type de flou mathématique précis est largement utilisé dans la vision artificielle artificiellement intelligente comme une étape de pré-traitement.

Par exemple, lorsqu'une photographie numérique est floue de cette manière, les bords des objets sont beaucoup plus faciles à détecter, ce qui permet à un ordinateur de les identifier automatiquement.

L'idée principale est la détection des structures les plus pertinentes de l'image à chaque échelle de la pyramide gaussienne en utilisant comme critère le seuil de détection en contraste.

3.2.2. Construction de la pyramide gaussienne :



Figure(2): Exemple d'une pyramide

La génération de la pyramide gaussienne est faite en commençant par une image initiale g_0 puis en passant par un filtre passe-bas (filtre de gauss) pour obtenir une image "réduite" g_1 .

L'image est "réduite" dans le sens où la densité spatiale et la résolution sont toutes deux diminuées.

Cette opération peut être poursuivie pour obtenir un ensemble d'images $\{g_0, g_1, g_2, \dots, g_n\}$ formant la structure de l'image pyramidale.

Image d'entrée $g_0(i, j)$:

L'image $g_1 = (g_0(i, j) * \text{filtreGauss})$, avec les dimensions i et j seront divisé par 2, jusqu'au g_7 (niveau 8).

❖ Le filtre de gauss est généré par la formule suivante :

$$1/2\pi\sigma^2 * \exp(-(x^2+y^2) / 2\sigma^2).$$

Le modèle d'Itti applique la pyramide gaussienne de 8 niveaux pour chaque caractéristique pour obtenir une pyramide gaussienne :

- Une pyramide d'intensité $I(\sigma)$: où σ représente la résolution, $\sigma \in [0 ; 8]$.
- 02 pyramides de couleurs : $R(\sigma)/G(\sigma)$, $B(\sigma)/Y(\sigma)$ pour les 02 oppositions des couleurs: rouge/vert, bleu/jaune.
- 4 pyramides d'orientation : pour les 4 orientations $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$.

3.3. Différence centre-voisins :

Les neurones visuels typiques sont les plus sensibles dans une petite région de l'espace visuel (le centre), alors que les stimuli présentés dans une région antagoniste plus large et plus faible concentrique avec le centre (l'entoure ou voisin) inhibent la réponse neuronale. Une telle architecture, sensible aux discontinuités spatiales locales, est particulièrement bien adaptée à la détection d'emplacements qui se distinguent de leur entourage et constitue un principe de calcul général dans la rétine, le noyau genouillé latéral et le cortex visuel primaire.[22]

C'est la différence entre les valeurs à différents niveaux de la pyramide, tel que :

✚ Le centre est un pixel à l'échelle $c = \{2, 3, 4\}$

- ✚ Le contour est le même pixel à l'échelle $s = c + \delta$, $\delta \in \{3,4\}$.
- ✚ L'opérateur \ominus représente la soustraction des valeurs à deux niveaux différence d'une pyramide ; cette soustraction nécessite une interpolation de la carte $M(s)$ pour qu'elle puisse avoir la même taille que $M(c)$.

❖ L'Intensité :

Le premier ensemble des cartes de caractéristiques est obtenu par le contraste de l'intensité, qui est détectée chez les mammifères par les neurones sensitifs, soit à des centres sombres sur des voisins lumineux, ou à des centres lumineux sur des voisins sombres. Les deux types de sensibilités sont simultanément calculé en un ensemble de six cartes.

Les valeurs des niveaux de résolution plus fine c sont soustraites aux valeurs des niveaux de résolution plus grossière. Alors la pyramide gaussienne d'intensité est générée $I(\sigma)$, la différence centre voisin est comme suit :

$$I(c,s) = |I(c) \ominus I(s)|$$

On obtient 06 cartes d'intensité:

1. $I(2,2+3) == I(2,5)$
2. $I(2,2+4) == I(2,6)$
3. $I(3,3+3) == I(3,6)$
4. $I(3,3+4) == I(3,7)$
5. $I(4,4+3) == I(4,7)$
6. $I(4,4+4) == I(4,8)$

❖ La Couleur :

Le deuxième ensemble des cartes de caractéristiques est construit de façon similaire pour les canaux de couleurs, qui sont représentés dans le cortex à l'aide d'un système appelé "color double-opponent" : dans le centre de leur champ récepteur, les neurones sont excités par une couleur (ex rouge) et inhibés par une autre (ex vert) le contraire est vrai pour les voisins. Les adversaires spatiaux et chromatiques existent dans le cortex visuel humain pour : rouge/vert, vert/rouge, bleu/jaune, jaune/bleu, donc les cartes $RG(c,s)$ et $BY(c,s)$ sont créés.

Les pyramides gaussiennes de couleurs sont générées $R(\sigma)$, $G(\sigma)$, $B(\sigma)$, $Y(\sigma)$, la différence centre voisin est comme suit :

02 ensembles de couleurs RV, BY

$$RV(c,s) = |(R(c)-G(c))\Theta (G(s)-R(s))|$$

$$BY(c,s) = |(B(c)-Y (c))\Theta(Y (s)-B(s))|$$

Après l'application de différence centre voisin on obtient 12 cartes de couleurs.

❖ Orientations :

L'information d'orientation locale est obtenue à partir des pyramides de Gabor orientées O (s, q), où s $[0..8]$ représente l'échelle et q $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ est l'orientation préférée .

(Les filtres de Gabor, qui sont le produit d'un réseau de cosinus et d'une enveloppe gaussienne 2D, se rapprochent du profil de sensibilité au champ réceptif (réponse impulsionnelle) des neurones sélectifs d'orientation dans le cortex visuel primaire [12]. s, q), coder, en tant que groupe, le contraste d'orientation local entre les échelles centrale et environnante ;

L'orientation est exprimée en degré se fait par : [22]

$$O(c,s,\theta) = |O(c,\theta)\Theta O(s,\theta)|$$

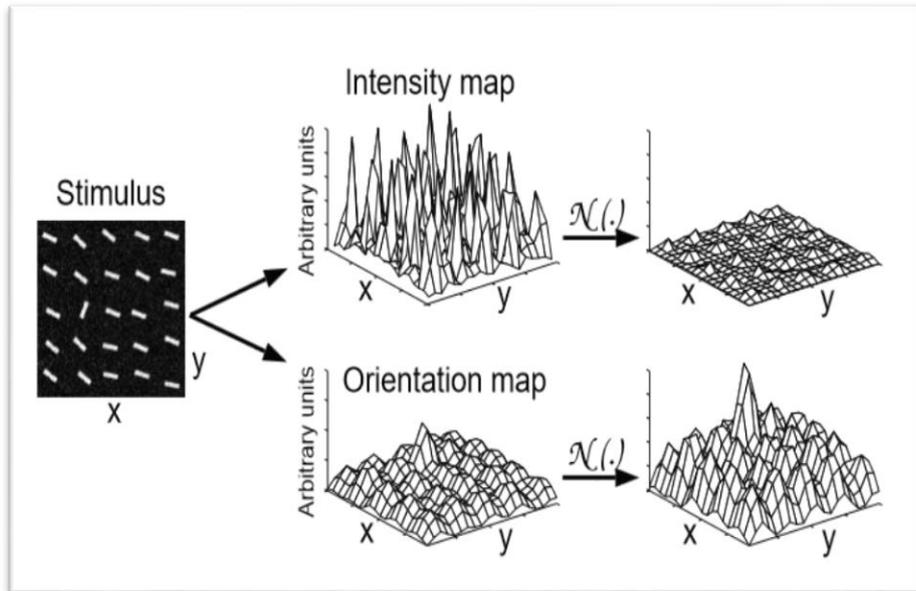
3.4. La normalisation :

Les cartes de caractéristiques seront sommées en vue de la création de la carte de saillance.

Alors que la dynamique de ces cartes peut être différentes car elles proviennent de différentes caractéristiques, il est nécessaire d'avoir une normalisation. De plus, cette normalisation renforce les cartes de caractéristiques qui ont un petit nombre de pics et diminue celles qui ont beaucoup de pics équivalents.

Ainsi, la normalisation de chaque carte de caractéristiques est effectuée de la manière suivante:

- ✚ Normaliser chaque pixel entre $[0, M]$. Ainsi, la valeur maximale globale de chaque carte est M .
- ✚ Calculer la valeur moyenne des maxima locaux m .
- ✚ Multiplier la carte par $(M - m)^2$. [13]



Figure(3) : normalisation des cartes

3.5. Combinaison des cartes :

Une carte de visibilité (conspicuity) est obtenue par la combinaison des cartes de caractéristiques après la normalisation à l'échelle 4 et pour chaque caractéristique.

Par addition point à point on obtient les résultats comme suit :

$$\text{Intensité} \quad I = \sum N(c,s).$$

$$\text{Couleurs} \quad C = \sum (N(RG(c,s)) + N(BY(c,s)))$$

$$\text{Orientation} \quad O = \sum N(\sum (N(O(c, s, \theta))))$$

3.6. Combinaison linéaire des cartes de visibilité (carte de saillance) :

Le but de la carte de saillance est de représenter la visibilité ou « saillance » à chaque endroit du champ visuel par une quantité scalaire et de guider la sélection des lieux fréquentés, en fonction de la distribution spatiale de la saillance. Une combinaison des cartes de caractéristiques fournit une entrée ascendante à la carte de saillance, modélisée comme un réseau neuronal dynamique.

Une difficulté dans la combinaison de différentes cartes de caractéristiques est qu'elles représentent des modalités a priori non comparables, avec des gammes dynamiques et des

mécanismes d'extraction différents. De plus, comme les 42 cartes d'entités sont combinées, les objets saillants apparaissant fortement dans seulement quelques cartes peuvent être masqués par le bruit ou par des objets moins saillants présents dans un plus grand nombre de cartes.

La carte de saillance finale (S) est construite en fusionnant les trois cartes de visibilité :

$$S = 1/3 (N(I) + N(C) + N(O))$$



Image d'entrée



Carte de saillance

Figure(4): carte de saillance finale

4. Conclusion :

Dans ce chapitre nous avons expliqué chaque étape du modèle de perception visuelle «Modèle Itti » pour extraire les caractéristiques visuelles dans le but de construire une carte de saillance. Nous remarquons que le modèle proposé, essaye d'extraire des caractéristiques de l'image initiale sous forme de d'images à niveaux de gris, qu'il combine ensuite en une carte de saillance, ce qui permettra de prendre des décisions concernant les traitements à entreprendre par la suite, qui se feront selon l'importance perceptuelle des zones de l'image initiale.

Le suivant chapitre concerne l'implémentation de ce modèle.

Chapitre 04

**Utilisation du modèle d'Itti
pour l'amélioration des
images**

(Implémentation et résultats)

1. Introduction :

Nous avons vu dans les chapitres précédents les différentes définitions des concepts de notre travail puis nous avons détaillé le modèle perceptuel que nous allons utiliser dans notre projet, qui est le modèle d'Itti, vu qu'il est simple et très utilisé dans la littérature.

Dans ce chapitre nous allons implémenter ce modèle, pour cela nous allons d'une part expliquer les différentes étapes d'implémentation de l'application et d'autre part exposer les résultats obtenus.

2. Langages de programmation :

- C++ :



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. [19]

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.

3. Bibliothèques utilisées :

- **OpenCV :**

OpenCV (Open Source Computer Vision) est une bibliothèque de fonctions de programmation principalement destinées à la vision par ordinateur en temps réel, tel que le traitement d'image. Initialement développé par Intel.

Nous avons utilisé la version Opencv 3.0 elle est compatible avec Visual Studio C++ 2015.

4. Outils de developement :

- **Visual Studio 2015 :**



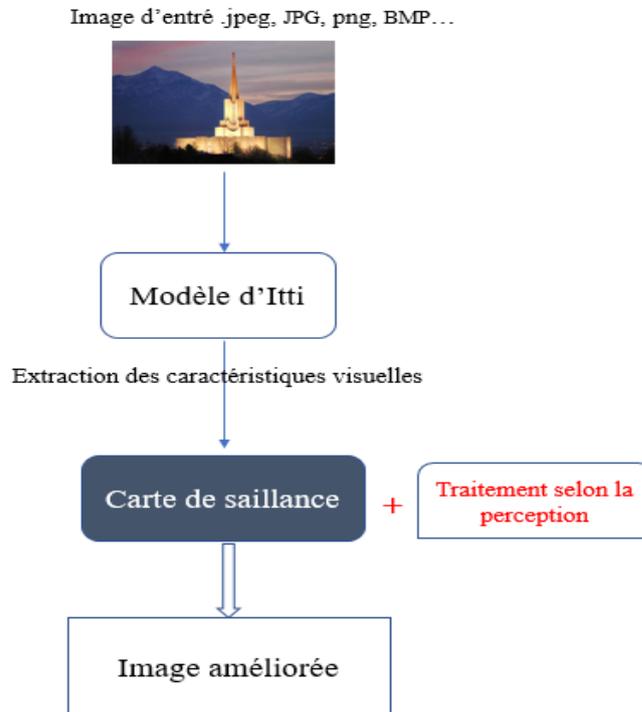
Visual Studio est un ensemble complet d'outils de développement qui permettant de générer des applications web ASP.NET, des services web XML, des applications bureautiques et des applications mobiles.

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 langage.

On a utilisé la version Visual Studio 2015.

5. Modèle d'Itti pour l'amélioration des images :

5.1. Schéma globale :



Figure(1): schéma globale du notre système

On a utilisé le modèle d'Itti pour l'extraction de la carte de saillance qui sélectionne les zones attirantes dans l'image qui sont captées par l'œil humain,

5.2. Amélioration du contraste :

La plage tonale d'une image représente le niveau de **contraste** de l'image. Elle est déterminée par la répartition des pixels dans l'image, des plus foncés (noirs) aux plus clairs (blancs).

On va améliorer l'image initiale, en changeant les pixels de l'image selon les pixels sélectionnées par la carte de saillance selon le contraste.

Le contraste est défini par la formule suivante :

$$g(i,j)=\alpha \cdot f(i,j)+\beta$$

Tel que $\alpha > 0$ et β ces paramètres représentent le gain et l'alignement, ils sont utilisés pour le contrôle du contraste et de la luminosité.

$$\alpha = (\text{Imax} - \text{Imin}) / (\text{Imax} + \text{Imin})$$

Imax = Intensité maximale (zone blanche)

Imin = Intensité minimale (zone noire)

$f(i,j)$: c'est l'image initial en entrée, c'est une matrice des pixels et chaque pixel on a 03 composantes des couleurs (rouge, vert et bleu).

$g(i,j)$: est l'image de sortie (image améliorée).

5.3. Les procédures et fonctions utilisées :

➤ Structure de données utilisées :

Nous avons utilisé les structures suivantes :

- ❖ L'image est déclarée comme une matrice 2d.
- ❖ Les cartes des couleurs, intensité et orientation sont représentées comme des matrices.

Mat imageI : représente la matrice d'intensité

- ❖ Les pyramides des caractéristiques sont représentées sous forme des vecteurs des caractéristiques, Par exemple ;
 - `vector<Mat> pyramidI(NUM_SCALES)`; c'est un vecteurs des image et désigne un pyramide d'intensité, tel que NUM_SCALES :est le nombre de niveau du pyramide
 - `vector<Mat> pyramidRG(NUM_SCALES)`; c'est le vecteur des image et désigne le pyramide de couleurs Rouge/Vert,
 - `vector<Mat> pyramid000(NUM_SCALES)`; représente un vecteur des image et désigne le pyramide d'orientation 0°
 - `vector<Mat_<float> > colors`; c'est un vecteurs des couleurs.
 - `vector<Mat> cspyrI`: vecteur de caractéristiques résulte de la différence centre-voisin de pyramide d'intensité

a. La 1ère étape : l'extraction des caractéristiques :

On va extraire les trois caractéristiques : intensité, couleurs et orientation

L'utilisation de la perception visuelle pour l'amélioration des images

➤ **L'intensité :**

Est extraite par la somme des trois couleurs composantes : rouge, vert et bleu/3.

➤ **Les couleurs :**

La fonction `split(scaled, colors);` permet d'extraire les couleurs dans l'image

`colors[0]` : est la couleurs rouge

`colors[1]` : est la couleurs vert

`colors[2]` : est la couleurs bleu

Puis on applique les formules suivantes pour avoir les 04 couleurs :

$$\text{Mat R} = r - (g + b) / 2;$$

$$\text{Mat G} = g - (b + r) / 2;$$

$$\text{Mat B} = b - (r + g) / 2;$$

$$\text{Mat Y} = (r + g) / 2 - \text{abs}(r - g) / 2 - b;$$

➤ **L'orientation :**

Selon la carte d'intensité on va extraire les cartes d'orientation en utilisant le filtre 2D de Gabor :

b. La 2ème étape : La différence centre voisin :

Pour chaque pyramide de caractéristique, nous appliquons la fonction différence centre voisin :

```
Mat operateCenterSurround(const Mat& center, const Mat& surround)
{
    Mat cmap(center.size(), center.type());
    resize(surround, cmap, cmap.size());
    cmap = abs(cmap - center);
    return cmap;}

```

La fonction `operatecenterSurrond` : permet de faire la différence point à point des images des niveaux des pyramides.

La fonction `resize()` : permet d'agrandir l'image voisine(`surround`) pour qu'elle ait la même taille que l'image centrale parce que les images de la pyramide n'ont pas la même taille

On sait que `voisin=centre + delta`, `centre = {2,3,4}`, `delta = {3,4}` en total on a 6 itérations

Tel que les nombres `{ 1,2,...8}` sont les niveau du pyramide gaussienne

```
vector<Mat> buildCenterSurroundPyramid(const vector<Mat>& pyramid)
{
    vector<Mat> cspyr(6);
    cspyr[0] = operateCenterSurround(pyramid[2], pyramid[5]);
    cspyr[1] = operateCenterSurround(pyramid[2], pyramid[6]);
    cspyr[2] = operateCenterSurround(pyramid[3], pyramid[6]);
    cspyr[3] = operateCenterSurround(pyramid[3], pyramid[7]);
    cspyr[4] = operateCenterSurround(pyramid[4], pyramid[7]);
    cspyr[5] = operateCenterSurround(pyramid[4], pyramid[8]);
    return cspyr;
}
```

c. 3ème étape : La normalisation :

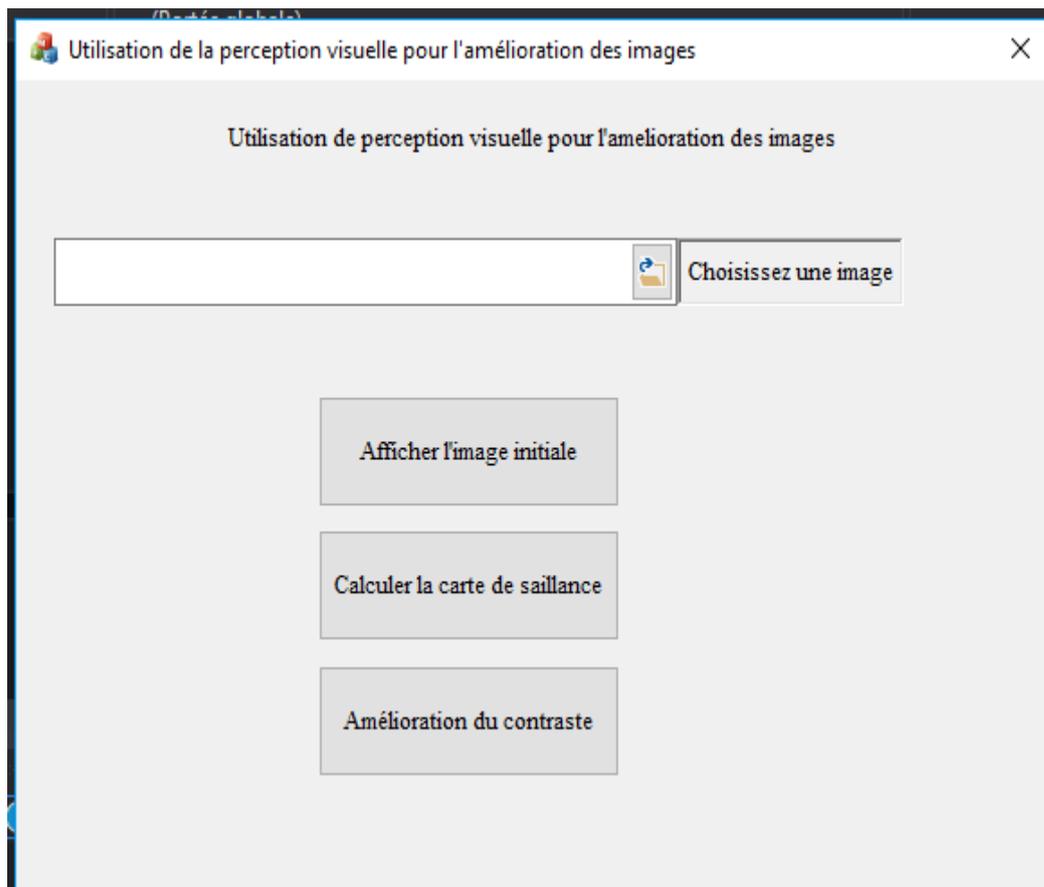
void normalizeRange(Mat& image) : procédure de normalisation qui permet de Normaliser la plage dynamique de l'image à [0,1]

void trimPeaks(Mat& image, int step): procédure de normalisation N: permet de l'accentuation unique du pic et suppression de plusieurs pics.

6. Résultats et Représentation des interfaces :

6.1. Interface d'application Principale :

Notre application est créée en C++ sous le « Visual Studio » relié par les bibliothèques OpenCV et MFC.

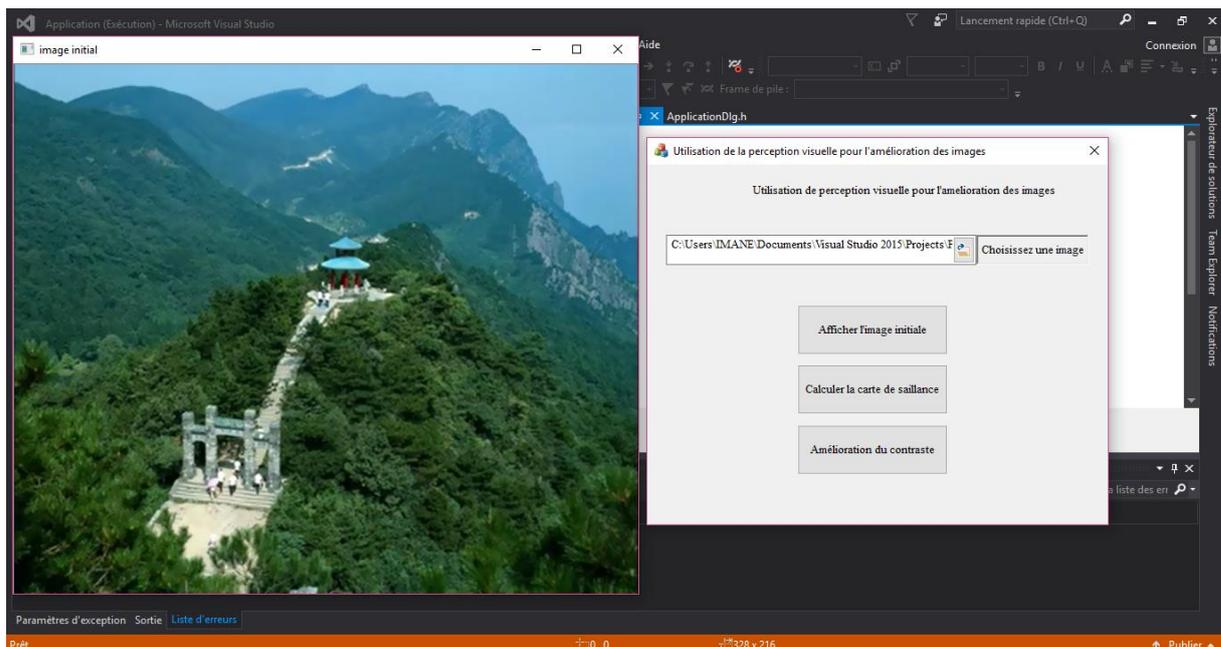


Figure(2) : interface

6.1.1. Charger une image :

En entrée nous avons une image RGB, à laquelle nous allons appliquer les étapes du modèle d'Itti pour extraire la carte de saillance, cette image doit être choisie depuis l'ordinateur à l'aide de champs choisissez une image.

Chapitre 04 : Utilisation du modèle d'Itti pour l'amélioration des images (Implémentation et résultats)



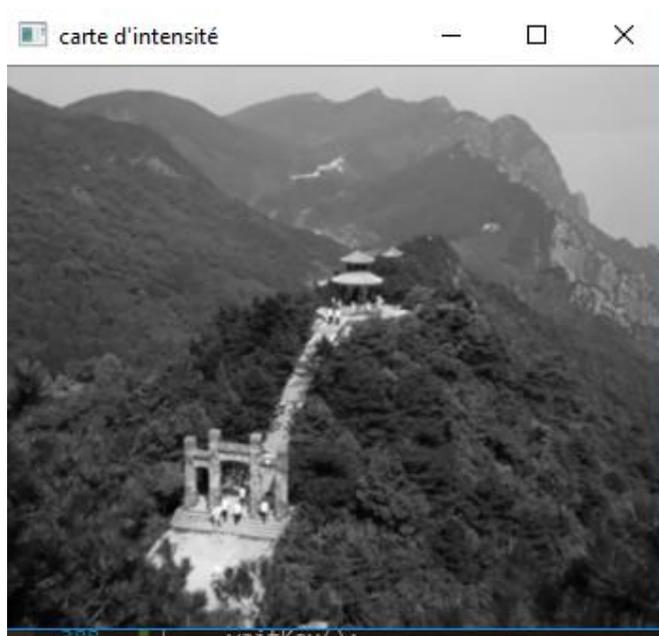
Figure(3) : Image d'entrée

Le deuxième bouton affiche la carte de saillance du modèle d'Itti de l'image ce qui affiche les parties les plus saillants dans cette image.

On passe par des différentes étapes pour obtenir cette carte de saillance :

- **Etape 1** : c'est l'extraction des caractéristiques visuelles ; intensité, couleurs et orientation
- **Intensité** : On obtient la carte d'intensité comme suit :

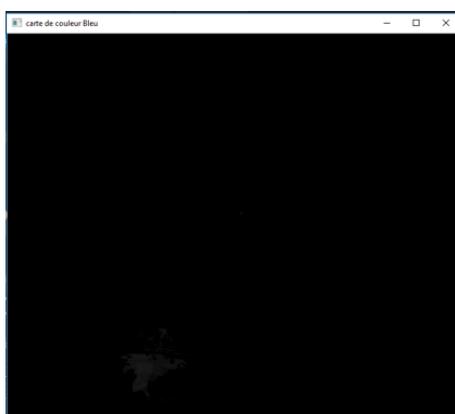
L'utilisation de la perception visuelle pour l'amélioration des images



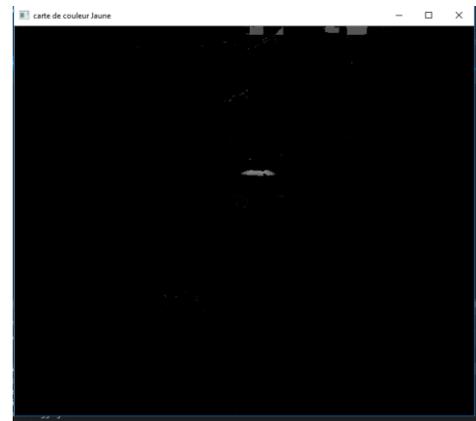
Figure(4) : Carte d'intensité

- **Couleurs :**

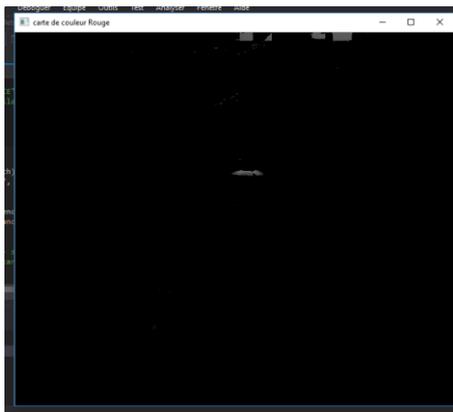
On obtient les cartes suivantes :



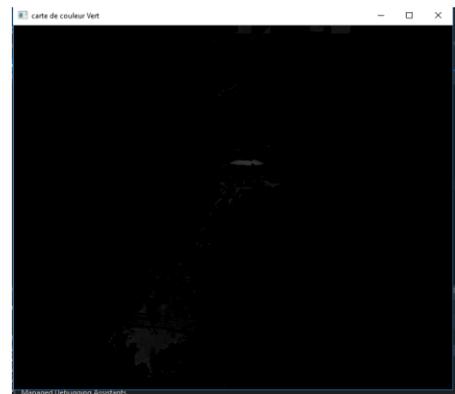
couleur «bleu »



couleur «jaune »



couleur «rouge »



couleur «vert »

Figure(5) : cartes des couleurs

- Orientation :

On obtient les cartes suivantes :



carte d'orientation 0°



carte d'orientation 45°



carte d'orientation 90°



carte d'orientation 135°

Figure(6) : les cartes d'orientation

- **Les pyramides gaussiennes :**

Chaque caractéristique on va appliquer la pyramide gaussienne

• **Pyramide d'intensité :**

On obtient 8 niveaux à chaque niveau c'est une image d'intensité, chaque image est 1/2 de l'image précédente.

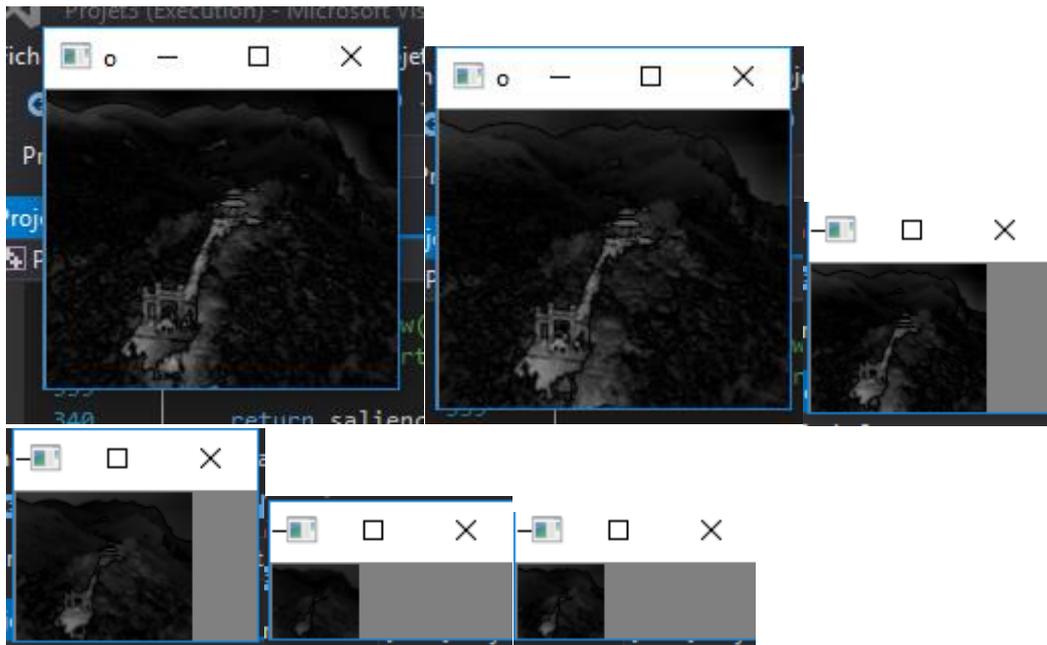


Figure(7) : Pyramide d'intensité

➤ **Etape 02 :** la différence centre _voisin :

Pour chaque caractéristique on obtient un ensemble de cartes.

• **Carte de caractéristiques d'intensité :** On obtient 6 cartes.

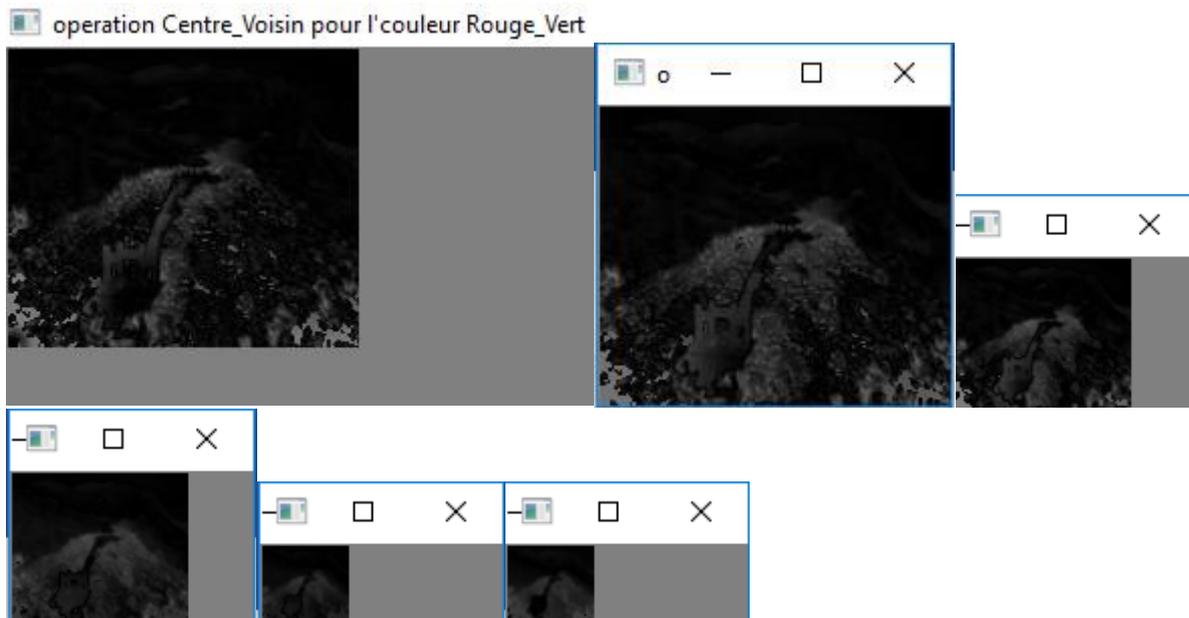


Figure(8) : carte de caractéristique d'intensité

- **Carte de caractéristiques de couleurs :**

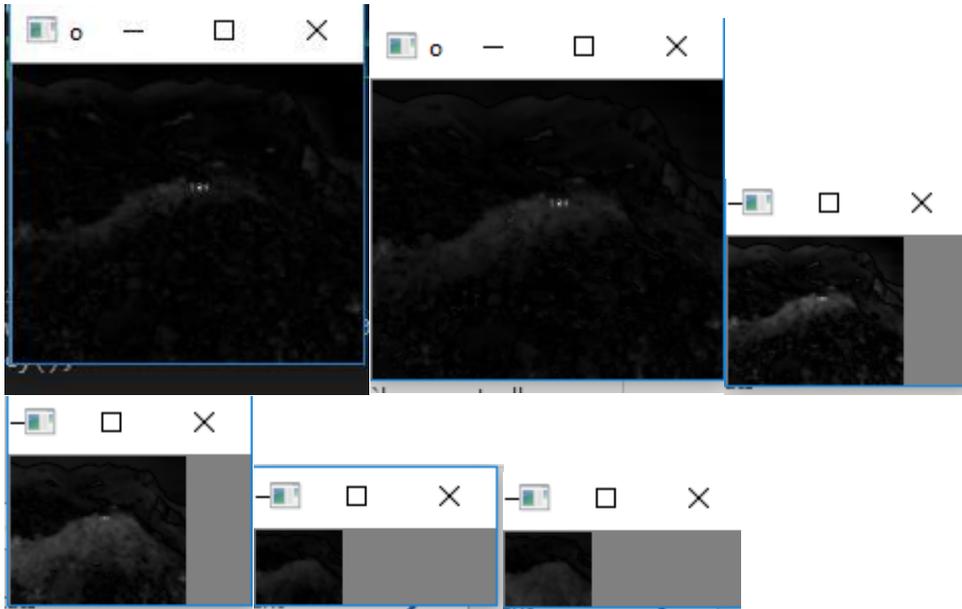
On obtient 12 carte ; 06 pour couleurs rouge_vert et 06 pour les carte Bleu_jaune

- **Couleur R-V : rouge_vert :**



Figure(9) : carte de caractéristique de couleur r_v

- **Couleur B_J :Bleu_jaune :**



Figure(10) : carte de caractéristique de couleur b_j

- **Orientation :**

Pour l'orientation on obtient 24 cartes, 06 cartes pour chaque orientation.

➤ **L'étape 03 :**

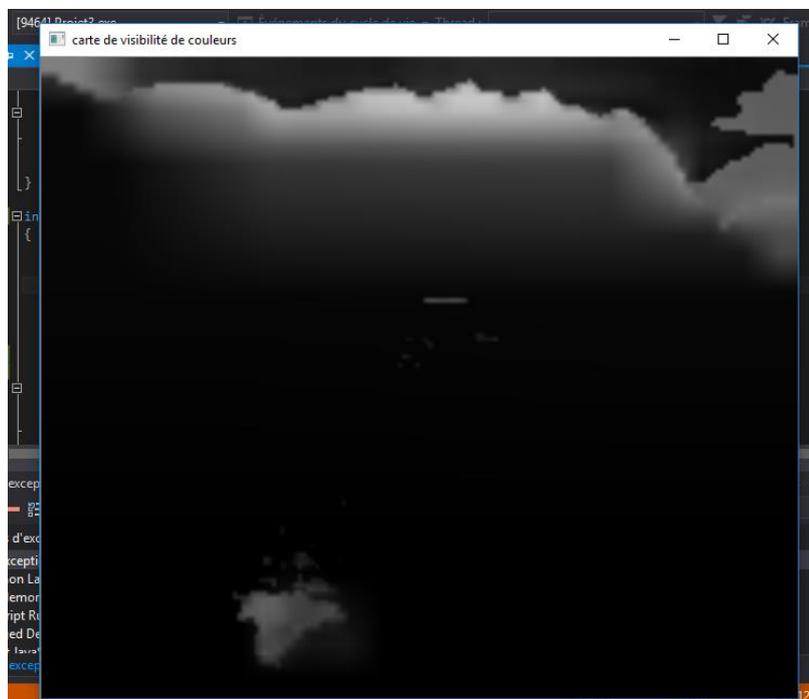
C'est la combinaison des cartes de caractéristiques après la normalisation pour obtenir les cartes de visibilité.

- **Carte de visibilité d'intensité :** c'est la combinaison des 06 cartes d'intensité obtenus de la différence centre -voisin.



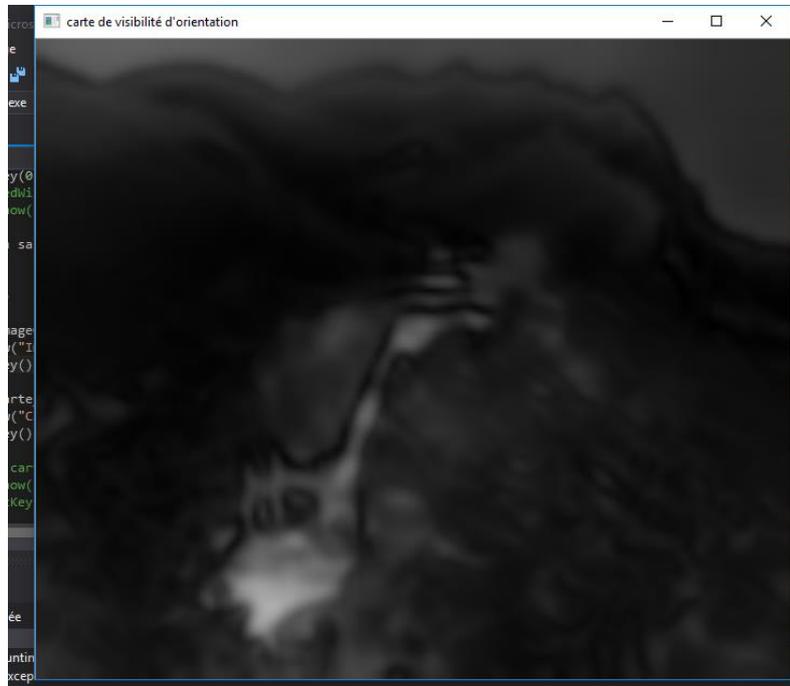
Figure(11) : carte de visibilité d'intensité

- **Carte de visibilité de couleurs** : c'est la combinaison des 12 cartes de couleurs obtenu de la différence centre -voisin.



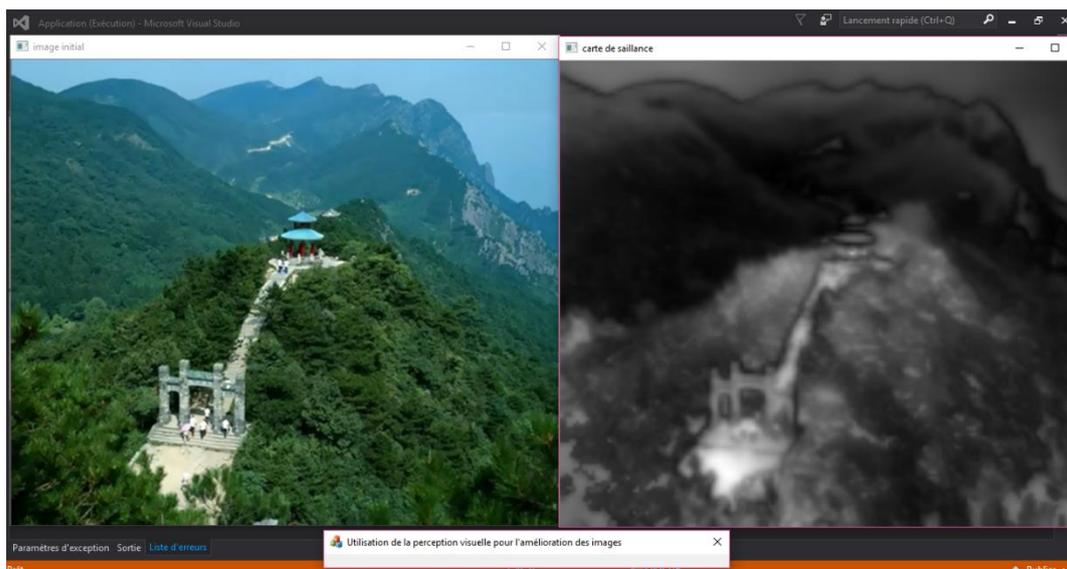
Figure(12) : carte de visibilité de couleur

- **Carte de visibilité d'orientation** : c'est la combinaison des 24 cartes d'orientations obtenu de la différence centre -voisin.



Figure(13) : carte de visibilité d'orientation

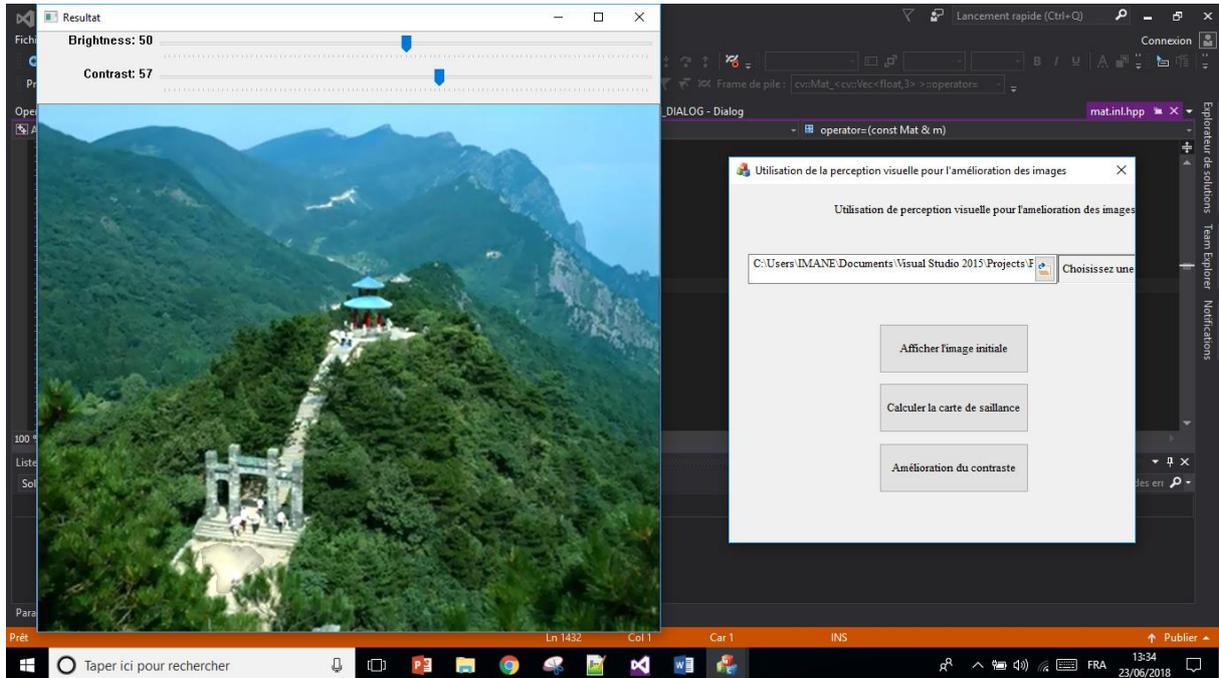
- **La dernière étape** : c'est la combinaison des 03 cartes de visibilité et comme ça on obtient la carte de saillance.



Figure(14) : carte de saillance

Amélioration de l'image :

Le dernier bouton c'est l'amélioration du contraste selon les parties sélectionnées par la carte de saillance, pour cela nous allons accentuer le contraste sur ces zones saillant qui apparaissent sur la carte de saillance et nous laissons les autres parties non accentuées.

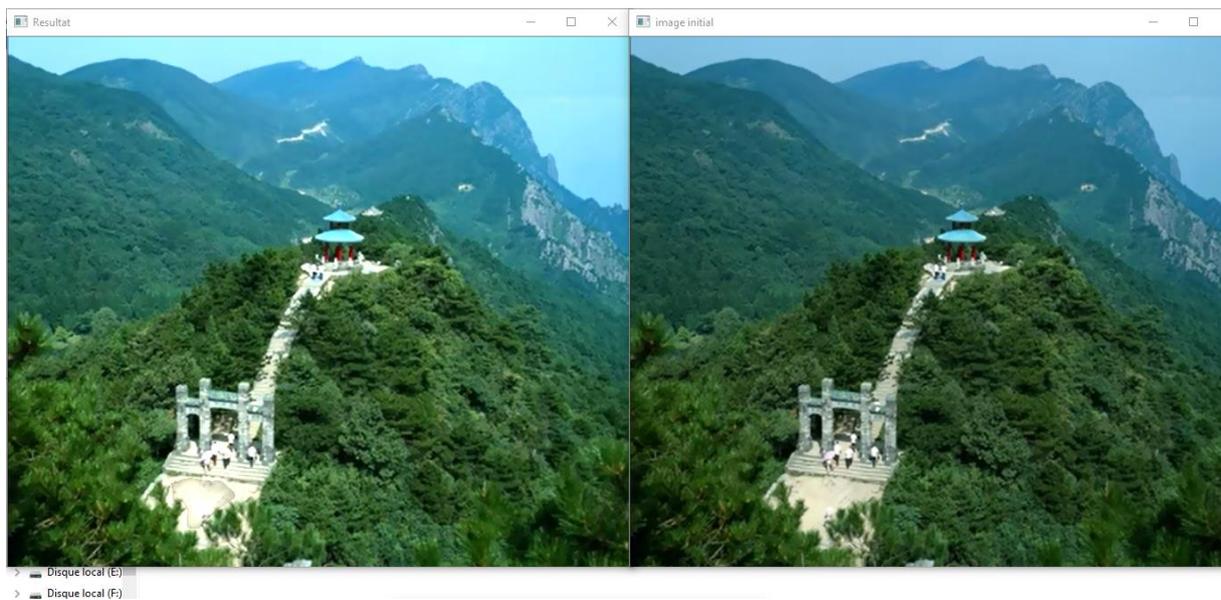


. *Figure(15) : accenturation du contraste et la luminosité sur l'image*

On a ajouté les barres en haut, c'est pour contrôler le contraste et la luminosité de l'image ils varient entre 0 et 100 lorsqu'on change la position de la barre le contraste change sur les zones sélectionnées par la carte de saillance.

Vers le 0 les parties sont sombres et vers le 100 l'image soit claire.

La figure 16 montre le résultat final qui affiche les parties saillantes de l'image, qui sont très contrastés que les autres parties non-saillants.



. Figure(16) : resultat finale

Conclusion générale :

Dans notre mémoire nous avons proposé la mise en œuvre d'un modèle de perception visuelle pour l'amélioration des images, c'est le modèle d'Itti qui permet dans une première phase d'extraire les caractéristiques visuelles pertinentes et les combiner dans une seconde phase pour créer une carte de saillance, qui affiche et sélectionne les zones importantes qui attire l'attention visuelle humain.

Nous avons obtenu les résultats suivants :

- ✓ Selon le modèle d'Itti, nous avons décomposé l'image d'entrée en trois caractéristiques visuelles qui sont : l'intensité, la couleur et l'orientation.
- ✓ Nous avons appliqué la pyramide gaussienne pour chaque caractéristique.
- ✓ Puis nous avons appliqué la différence centre voisin entre les niveaux des pyramides de chaque caractéristique.
- ✓ Ensuite, nous avons combiné les cartes obtenues par la différence centre voisin, ce qui a engendré les trois cartes de visibilité.
- ✓ Enfin, nous avons combiné les cartes de visibilité, ce qui nous a permis d'obtenir la carte de saillance, qui est une image 2D représentant l'attention visuelle.
- ✓ Après avoir construit la carte de saillance, nous avons appliqué comme exemple de validation, un traitement consistant en l'accentuation le contraste et de la luminosité sur les partis sélectionnés par cette carte.

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.

- ✓ **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. Les modèles détaillés dans ce projet ne concernent que les scènes statiques (les images) qui ne considèrent que des caractéristiques statiques dans l'image : luminance, couleur et orientation, mais ils

existent aussi des travaux qui se proposent d'étudier les scènes dynamiques (vidéos) en ajoutant deux caractéristiques aux trois précédents : le mouvement et le flicker (variation temporelle de l'intensité). Il serait intéressant de faire des études approfondies dans ce contexte, et de proposer des modèles de détection de saillance 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.
- ✓ **Application des différents filtres de traitement d'images**, afin d'améliorer le contenu de l'image de façon sélective selon la perception visuelle, ce qui représente un bon compromis entre qualité de l'image obtenue et le temps de calcul.
- ✓ **Tracking** : le suivi des objets.

Bibliographie et webographie :

[1] <https://www.guide-gestion-des-couleurs.com/oeil-perception-couleurs.html>

consulté le 03/11/2017

[2] https://www.assistancescolaire.com/eleve/1L/ens_sciences/reviser-le-cours/la-perception-visuelle-de-la-retine-au-cerveau-1_w101 consulté le 03/11/2017 à 20:51

[3] <http://www.intellego.fr/soutien-scolaire--/aide-scolaire-svt/cone-et-batonnet-photorecepteurs-de-la-retine-/50166> consulté le 03/03/2018 à 21:50

[4] <https://fr.wikipedia.org/wiki/Rétine> consulté le 03/11/2017 à 22:13

[5] <https://www.youtube.com/watch?v=2lIRHP--ROY> de l'œil au cerveau construction de la perception visuelle. consulté le 08/12/2017 à 23 :23

[6] Morgan Veyret « Focalisation de l'attention visuelle » ,6 janvier 2004

[7] : <https://askabiologist.asu.edu/b%C3%A2tonnets-et-c%C3%B4nes> consulté le 04/04/2018 à 12 :34

[8] <http://la-perception-visuelle.e-monsite.com/medias/images/vision-cerveau-traduit.jpg> consulté le 12/02/2018 à 13:01

[9] : <http://www.profil-couleur.com/lc/018-couleur-definition.php> consulté le 05/12/2017

[10]

https://www.unine.ch/files/live/sites/physique/files/TP/Exp%C3%A9riences_avec_FM/Exp12-Polarisation_de_la_lumiere.pdf consulté le le 05/01/2018 à 20 :42

[11] : Eric Marchand. Colorimétrie. Cours. INRIA. Université de Rennes 1. 2017.

[12] : Antoine Manzanera. Traitement Et Reconnaissance d'Images. Cours TERI – Master IAD UPMC Paris 6. 2017.

[13] : Aurélie Bugeau. ATTENTION VISUELLE MULTI-ECHELLE, mémoire de stage, DEA Imagerie Vision Robotique. INRIA Rhône-Alpes. 2004.

L'utilisation de la perception visuelle pour l'amélioration des images

[14] : Taha RIDENE. Mécanismes d'Attention Visuelle sur Rétine Artificielle, Mémoire de stage de 2ème année Master IPCC spécialité Image ; Université Paris V, René Descartes. 2006.

[15] : Matthieu Perreira da Silva « Modèle computationnel d'attention pour la vision adaptative », Thèse de doctorat. Université de La Rochelle. 2010.

[16] : Ilyes Aboulaziz « L'étude de la saillance visuelle dans le contexte de la catégorisation des scènes naturelles », Projet de fin d'études, master informatique et télécommunication, Université du rabat. 2010.

[17] : Tien Ho Phuoc « Développement et mise en œuvre de modèles d'attention visuelle » Thèse de doctorat. Université de Grenoble. 2010.

[18] : Matthieu Perreira da Silva, Vincent Courboulay « Implémentation et évaluation d'un modèle d'attention pour la vision adaptative » Traitement du Signal, Lavoisier, 2012, 28 (6), pp.611-641. 2012

[19]: A. Treisman : "A Feature-Integration Theory of Attention" Université de Bretagne Columbia and G. Gelade Oxford University (1980)

[20] : http://theses.univ-lyon2.fr/documents/getpart.php?id=lyon2.2007.keita_1&part=199086; consulté le 30-04-2018 à 20 :00

[21] : Guillaume Rousselet : « Catégorisation visuelle rapide des scènes naturelles : limites du parallélisme et spécificité des visages. Une étude comportementale et électrophysiologique chez l'humain ». Thèse de doctorat. Ecole des Hautes Etudes en Sciences Sociales (EHESS), 2003.

[22] : Laurent Itti, Christof Koch, and Ernst Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis" IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 20 Issue 11, November 1998. Pages 1254-1259

[23]: http://theses.univ-lyon2.fr/documents/getpart.php?id=lyon2.2007.keita_1&part=199089 ; consulté le 02/05/2018