



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

Université Saad Dahlab, Blida  
USDB.

Faculté Des Sciences  
Département Informatique



Mémoire Pour L'obtention  
Du Diplôme D'ingénieur D'état En Génie Informatique  
Option : Intelligence Artificielle

Sujet :

**Les descripteurs couleurs MPEG-7  
pour l'Indexation et Recherche dans  
une base d'images**

Présenté par :  
LAIEB Mustapha  
MOSSAAB Ali

proposé et dirigé par :  
Mlle SOUAMI Feryel

Soutenu le: *27/11/2008*, devant le jury composé de :

Nom président du jury : *BENBLIDJA NADJIA*

Nom examinateur 1 : *Cherif ZAHAR*

Nom examinateur 2 : *NADOL Malek*

MIG-004-215-1

## *Remerciements*

En premier lieu, nous remercions *Dieu* le tout puissant qui nous a donné la santé, la volonté, la patience, ainsi que le courage pour étudier et aller jusqu'au bout.

Notre profonde gratitude et nos sincères remerciements vont à notre promotrice *Dr SOUAMI Feryel* pour avoir proposé ce sujet, et de nous avoir dirigé et encouragé tout au long de la réalisation de ce travail.

Nous remercions encore *Mr M.NADIL* qui nous a aidé à réaliser ce mémoire.

Nous remercions également les membres du jury :

- *Mme BENBLIDIA* pour avoir accepté de présider notre soutenance.
- *Mr CHERJF ZAHAR* et *Mr M.NADIL* pour avoir examiné notre thèse.

Merci enfin à tous ceux, famille et amis qui nous ont soutenu ou aidé durant ce projet.

*Ali et Mustapha*

## *Dédicaces*

*A celui dont l'absence n'exclut pas sa présence dans mon esprit, mon père.*

*A celle dont l'existence est ma raison de vivre et sans elle ma réussite n'aurait jamais lieu,...*

*Que Dieu te protège ma très chère mère.*

*A tous mes chers frères et sœurs, neveux et nièces.*

*A toute ma grande famille et surtout mes deux cousins «Hocine et Belkacem»*

*A tous mes chers amis :*

*Mohamed, Lamine, Mahmoud, Othmane, Nassim, Djamel, Azzi, Yacine, Belkas, Hichem, Amine, Toufik, ..*

*A tous mes camarades de l'université de Blida et surtout Younes, Smail et Ziane.*

*A mon binôme Mustapha et toute sa famille.*

*A tous ceux que j'ai eu le plaisir de connaître durant toute ma vie.*

*Je dédie ce modeste travail*

*Ali*

## *Dédicaces*

*A ceux dont l'absence n'exclut pas leur présence dans mon esprit et ceux dont l'existence est ma raison de vivre, mes parents.*

*A l'esprit de mes deux grandes mères.*

*A mes chères sœurs et frères surtout ma petite sœur Ilhem.*

*A mes chères tantes et oncles.*

*A toute ma grande famille et surtout mes deux cousins «Zoheir, Mohamed et Zaki»*

*A tous mes chers amis de Kouba : Hichem, Rafik, Amine, Hafid, Alaedine, Mokrane, Nazim, Ali, Samir, Sabine.*

*A tous mes amis de Damous.*

*A tous mes camarades de l'université de Blida et surtout Bachir, Younes, Hocine, Moussa, Badro, Ziane.*

*A mon binôme Ali et toute sa famille.*

*A tous ceux qui me sont chers.*

*Je dédie ce modeste travail  
Mustapha*

## RÉSUMÉ

La couleur est largement utilisée, elle est considérée comme un attribut robuste dans l'indexation par le contenu. Sa représentation se fait soit par une représentation fréquentielle, soit par le choix d'un espace couleur. La phase de distinction entre deux images est faite grâce au calcul de la distance entre le vecteur des caractéristiques de l'image requête et le vecteur des caractéristiques de l'image de base.

L'objectif de notre travail est de développer un système qui permet l'indexation et la recherche d'images par le contenu couleur. Notre système se base essentiellement sur les descripteurs de contenu couleur définis par le standard MPEG-7. Dans ce projet, nous construisons nos descripteurs en déterminant parmi eux lesquels sont les meilleurs pour la recherche dans une base d'images.

### **Mots-clés :**

Indexation et recherche d'images par le contenu.

Les descripteurs couleurs MPEG-7.

# SOMMAIRE



## *INTRODUCTION GÉNÉRALE*

### *Chapitre I : Image et couleur*

Introduction.....	1
I.1 Caractéristiques d'une image numérique.....	2
I.1-1 Pixel, connexité et niveaux de gris .....	2
I.1-2 Résolution.....	3
I.1-3 Luminance.....	3
I.1-3 Contraste.....	3
I.2 Description d'une image numérique.....	4
I.2-1 Histogramme .....	4
I.2-2 Contours et régions.....	4
I.2-3 Homogénéité.....	4
I.3 Formats d'images numériques.....	5
I.4 La couleur.....	6
I.4-1 Perception de la couleur.....	6
I.4-2 Caractéristiques de la couleur.....	6
I.4-2-1 Luminosité.....	7
I.4-2-2 Teinte.....	7
I.4-2-3 Saturation.....	7
I.4-2-4 Intensité.....	8
I.4-3 Codage de la couleur.....	8
I.5 Espaces couleurs.....	9
Conclusion.....	9

## ***Chapitre II : Indexation et Recherche d'images***

Introduction .....	10
II.1 Définition de l'indexation .....	10
II.2 Approches de l'indexation .....	11
II.2.1 Approche globale .....	11
Couleur.....	11
Texture.....	12
Forme .....	13
II.2.2 Approche spatiale .....	13
II.4 Recherche d'information .....	13
II.5 Architecture générale du système d'indexation et de recherche d'images .....	14
II.6 L'interrogation .....	15
II.7 Mesures de similarités .....	16
Conclusion .....	17

## ***Chapitre III : Les descripteurs couleurs MPEG\_7***

Introduction.....	18
III.1 Définition de MPEG_7 .....	18
III.2 Fonctionnalités majeurs du MPEG_7 .....	19
III.3 Portée de standard MPEG_7 .....	20
III.4 Les applications MPEG_7 .....	20
III.5 Structures de MPEG_7 .....	20
III.6 Les descripteurs visuels MPEG_7 .....	22
III.7 Les descripteurs couleurs MPEG_7 .....	23
III.7.1 Le descripteur Color Space.....	24
III.7.1.1 L'espace couleur RGB.....	24
III.7.1.2 L'espace couleur YCbCr.....	25
III.7.1.2 L'espace couleur HSV.....	25
III.7.1.2 L'espace couleur LUV.....	26
III.7.1.2 L'espace couleur HMMD.....	27

III.7.2 Le descripteur Dominant Color (DCD).....	28
III.7.3 Le descripteur Color Layout (CLD).....	33
III.7.4 Le descripteur Scalable Color (SCD).....	36
III.7.5 Le descripteur Group Of Pictures or Group Of Frames ( GOP/GOF).....	39
III.7.6 Le descripteur Color Structure (CSD).....	41
Conclusion .....	44

### ***Chapitre IV : Conception et mise en œuvre***

Introduction .....	45
IV.1 Architecture générale du système .....	46
IV.2 Description de l'architecture.....	47
IV.2.1 Module d'indexation .....	47
IV.2.1.1 Extraction de descripteur Color Layout.....	47
IV.2.1.2 Extraction de descripteur Scalable Color.....	49
IV.2.1.3 Extraction de descripteur Dominant Color.....	49
IV.2.2 Modèle de représentation des données.....	50
IV.2.3 Module de recherche.....	52
IV.2.3.1 Recherche unique.....	53
IV.2.3.2 Recherche multiple (combinaison des descripteurs).....	57
Conclusion .....	58

### ***Chapitre V : Réalisation et Résultats expérimentaux***

Introduction .....	59
V.1 Résultats d'indexation.....	60
V.1.1 Descripteur Color Layout.....	60
V.1.2 Descripteur Scalable Color.....	61
V.1.3 Descripteur Dominant Color.....	61
V.2 Critères d'évaluation.....	62

V.3 Résultats de la recherche.....	62
V.3.1 Descripteur Color Layout.....	63
V.3.2 Descripteur Scalable Color.....	64
V.3.3 Descripteur Dominant Color.....	65
V.3.4 Résultats de la combinaison.....	67
V.4 Étude des résultats .....	70
V.4.1 Descripteur Color Layout.....	71
V.4.2 Descripteur Scalable Color.....	71
V.4.3 Descripteur Dominant Color.....	72
V.4.4 Combinaison des descripteurs.....	73
Conclusion.....	74

***CONCLUSION GÉNÉRALE ET PERSPECTIVES***

***BIBLIOGRAPHIE***

## Liste des Figures

Figure 1.1 Représentation des deux types d'images.....	1
Figure 1.2 Lettre A affichée comme un groupe de pixels.....	2
Figure 1.3 Histogramme couleurs d'une image.....	3
Figure 1.4 Perception de la couleur.....	6
Figure 2.1 L'architecture d'indexation et de la recherche d'images par le contenu.....	14
Figure 3.1 Place du standard MPEG-7 dans le processus d'accès au contenu.....	20
Figure 3.2 Diagramme de hiérarchie de MPEG-7.....	21
Figure 3.3 Les différents descripteurs couleurs.....	23
Figure 3.4.a Représentation cube de l'espace couleur RGB.....	24
Figure 3.4.b Mode additif de RGB.....	24
Figure 3.5 Représentation des couleurs dans l'espace YCbCr.....	25
Figure 3.6 Représentation des couleurs dans l'espace HSV.....	26
Figure 3.7 L'espace de couleurs LUV.....	27
Figure 3.8 Représentation double cone de l'espace couleur HMMD.....	28
Figure 3.9 Exemple d'un descripteur Dominant Color représenté sous format XML.....	31
Figure 3.10 Le processus d'extraction de CLD.....	34
Figure 3.11 Descripteur Color Layout représenté sous format XML.....	35
Figure 3.12 Extraction du descripteur Scalable Color.....	37
Figure 3.13 Descripteur Scalable Color représenté sous format XML.....	39
Figure 3.14 Élément structurant pour le descripteur Color Structure.....	42
Figure 3.15 Élément structurant avec des images de différentes tailles.....	43
Figure 3.16 Représentation de descripteur Color Structure sous format XML.....	44
Figure 4.1 Architecture d'un système de recherche d'image par le contenu couleur.....	46
Figure 4.2 Module d'indexation.....	47
Figure 4.3 Processus d'extraction de CLD (étapes).....	48
Figure 4.4 Module de Recherche.....	52
Figure 5.1 Exemple d'images.....	59
Figure 5.2 Fenêtre de recherche.....	60
Figure 5.3 Résultat de la recherche en utilisant le descripteur Color Layout.....	63
Figure 5.4 Résultat de la recherche en utilisant le descripteur Scalable Color.....	64
Figure 5.5 Résultat de la recherche en utilisant le descripteur Dominant Color (DCD).....	58
Figure 5.6 Résultat de la recherche en combinant SCD et CLD (filtrage).....	58
Figure 5.7 Résultat de la recherche en combinant DCD et CLD (poids).....	59

## *Liste des Tableaux*

<b>Tableau 3.1</b> : Fonctionnalités majeurs de standard MPEG-7.....	19
<b>Tableau 3.2</b> : Représentation du descripteur Dominant Color.....	22
<b>Tableau 3.3</b> : Représentation du descripteur Color Layout.....	35
<b>Tableau 3.4</b> : Nombre du niveau H,S et V en fonction du nombre de coefficients.....	38
<b>Tableau 3.5</b> : Représentation du descripteur Scalable Color.....	38
<b>Tableau 3.6</b> : Représentation du descripteur GOP or GOF.....	41
<b>Tableau 3.7</b> : Quantification de l'espace HMMD pour le CSD.....	44
<b>Tableau 4.1</b> : Exemple de valeurs poids des coefficients CLD.....	55
<b>Tableau 5.1</b> : résultats de la recherche en utilisant Color Layout.....	71
<b>Tableau 5.2</b> : résultats de la recherche en utilisant Scalable Color.....	72
<b>Tableau 5.3</b> : résultats de la recherche en utilisant Dominant Color.....	73
<b>Tableau 5.4</b> : résultats de la recherche (comparaison des résultats).....	74

# *Introduction générale*



La technologie numérique a révolutionnée les moyens de communications avec l'arrivée entre autre des téléphones portables, de l'internet à haut débit et de la télévision numérique. L'explosion des moyens de communications a conduit à une augmentation spectaculaire des informations mise à disposition. L'information peut être du type texte, son, image fixe ou animée, ou une combinaison de ces différents types de données sous des formats multimédia.

Une des problématiques que pose cette diversité de l'information est son exploitation. Les données sont souvent stockées dans des bases de données multimédia et il s'agit alors pour l'utilisateur d'y retrouver le ou les documents qui l'intéressent, souvent de manière intuitive.

Cependant, rechercher dans une grande base de données des informations précises, en un temps raisonnable nécessite au préalable de structurer cette base de donnée. S'agissant de données de type alphanumériques (texte), des technologies et des méthodes sont déjà mises au point. Elles utilisent souvent les mots clés comme des descripteurs de contenu permettant son indexation. Pour les données de type multimédia, les méthodes nécessaires pour leur description sont encore en cours de développement.

Il s'agit de décrire le contenu de ces données afin de les indexer. Les données multimédia se composent de son, de vidéo, de texte, et d'images. Le son est une donnée particulière pour laquelle les travaux de la téléphonie et de la télévision ont permis de mettre en évidence des schémas de description amenant à des dispositifs d'indexation. La vidéo est elle en fait une succession d'images. Il s'agit donc de décrire, aussi bien pour la vidéo que pour les images, leur contenu en déterminant les éléments pertinents qui les décrivent. Ces éléments de description (descripteurs) seront alors les index de la base d'images ou vidéos.

L'indexation d'images constitue donc une des clés importantes pour répondre au problème de la consultation et la recherche d'information dans de grande base de données. Cependant, l'image comporte intrinsèquement une information riche et difficile à décrire. L'indexation et la recherche d'images par le contenu est un problème complexe car différents critères permettent sa description : la couleur, la forme et la texture.

Dans cette étude, nous nous sommes intéressés à la caractéristique couleur contenu dans des images. Le laboratoire LRIA (USTHB) ayant déjà mené des études précédentes sur le contenu couleur des images, s'est orienté comme nombre de laboratoire vers les descripteurs issus des standards tel MPEG7. Ce travail s'inscrit dans ce nouvel axe.

Afin de mener à bien ce travail, nous avons suivis différentes étapes, réparties dans les chapitres suivants.

- Dans le **premier chapitre**, nous donnerons quelques généralités sur la couleur et l'image numérique qui est l'objet principal de notre mémoire.
- Puis, dans le **deuxième chapitre**, nous étudions les principes de l'indexation et de la recherche d'images par le contenu.
- Ce travail ayant pour objectif d'étudier les descripteurs de la norme MPEG-7 et leur apport dans l'indexation, nous avons étudié la norme et ses composantes dans le **troisième chapitre**. Les descripteurs couleurs définis dans la norme MPEG-7 y sont décrits.
- Le **quatrième chapitre** portera sur notre conception d'un système d'indexation et de la recherche d'images basé sur des descripteurs couleurs de MPEG-7.
- Enfin, dans le **cinquième chapitre**, l'implémentation sera présentée. Nous y présenterons aussi les résultats expérimentaux avec une discussion de ces résultats obtenus par différents tests.
- Une conclusion et des perspectives concluront ce travail.

## INTRODUCTION

Contrairement aux images obtenues à l'aide d'un appareil photo, ou dessinées sur du papier, les images manipulées par un ordinateur sont numériques (représentées par une série de bits). L'image numérique est l'image dans la surface est divisée en éléments de tailles fixes appelés cellules ou *pixels* [1].

Il existe deux sortes d'images numériques : les images vectorielles, les images matricielles [2].

- ✓ Une *image vectorielle* est l'image dans laquelle les données sont représentées par des formes géométriques simples qui sont décrites d'un point de vue mathématique. Par exemple, un cercle est décrit par la position de son centre et son rayon. Ces images sont essentiellement utilisées pour réaliser des schémas ou des plans. Elles occupent peu de place en mémoire, ainsi, elles peuvent être redimensionnées sans pertes d'informations.
- ✓ Une *image matricielle* est formée d'une matrice (tableau) de points ou *pixels*. Par exemple, les images vues sur un écran de télévision ou une photographie sont des images matricielles. Contrairement aux images vectorielles, les images bitmap occupent un grand espace mémoire avec une durée de traitement plus grande.

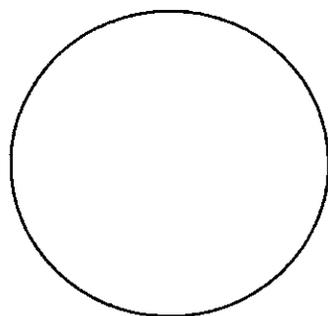


Image vectorielle

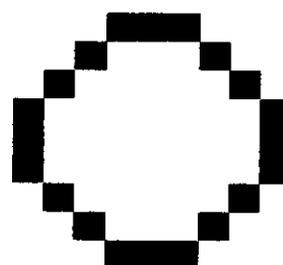


Image matricielle

Figure 1.1 : Représentation des deux types d'images.

## I.1 CARACTÉRISTIQUES D'UNE IMAGE NUMÉRIQUE

### I.1.1 Pixel, connexité et niveaux de gris

#### ➤ Pixel [3]

Le *pixel*, ou bien élément d'image (en anglais *picture element*) est le plus petit élément (l'unité de base) d'une image numérique.

Dans une image, à chaque pixel est attribuée une intensité lumineuse. La qualité d'une image numérique dépend du nombre de pixels. Ainsi, le nombre de bits nécessaires pour coder un pixel dépend du nombre de couleurs affichables, par exemple pour afficher 256 couleurs, le pixel sera codé sur 8 bits.

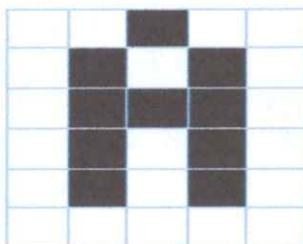
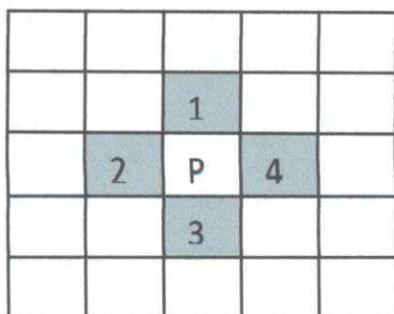


Figure1.2 : lettre A affichée comme un groupe de pixels.

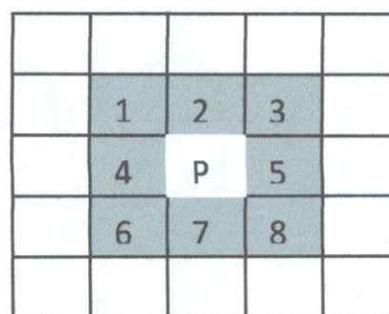
#### ➤ Connexité [4]

C'est le voisinage d'un pixel. On peut distinguer deux types de connexité :

- La 4-connexité : le voisinage d'un pixel  $P(i,j)$  est dit 4-connexe s'il est formé de l'ensemble de 4 pixels de coordonnées spatiales  $(i+1,j)$  et  $(i-1,j)$  et  $(i,j-1)$  et  $(i,j+1)$ .
- La 8-connexité : il est dit 8-connexe si en plus des 4 voisins, l'ensemble comprend quatre coordonnées spatiales.



Voisinage 4-connexe



voisinage 8-connexe

### ➤ Niveaux de gris [4]

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires.

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

### 1.1.2 Résolution [1]

La résolution définit le nombre de pixels par unité de longueur (centimètre ou pouce). La résolution d'une image numérique s'exprime en PPI (Pixel Per Inch) ou PPP (Pixels Par Pouce).

### 1.1.3 Luminance [1]

C'est le degré de luminosité des points de l'image. Elle est définie aussi comme étant le quotient de l'intensité lumineuse d'une surface par l'aire apparente de cette surface.

Une bonne luminance se caractérise par :

- Des images lumineuses (brillantes).
- Un bon contraste.
- L'absence de parasites.

### 1.1.4 Contraste [1]

C'est l'opposition entre les régions sombres et les régions claires d'une image, Le contraste est défini en fonction des luminances de deux zones d'images. Si  $L_1$  et  $L_2$  sont les degrés de luminosité respectivement de deux zones voisines  $Z_1$  et  $Z_2$  d'une

image, le contraste  $C$  est défini par le rapport : 
$$C = \frac{L_1 - L_2}{L_1 + L_2}$$

## I.2 DESCRIPTION D'UNE IMAGE NUMÉRIQUE

### I.2.1 Histogramme [8]

Un histogramme est un graphique statistique permettant de représenter le nombre de pixels pour chaque intensité lumineuse.

Pour diminuer l'erreur de quantification, pour comparer deux images obtenues sous des éclairages différents, ou encore pour mesurer certaines propriétés sur une image, on modifie souvent l'histogramme correspondant.

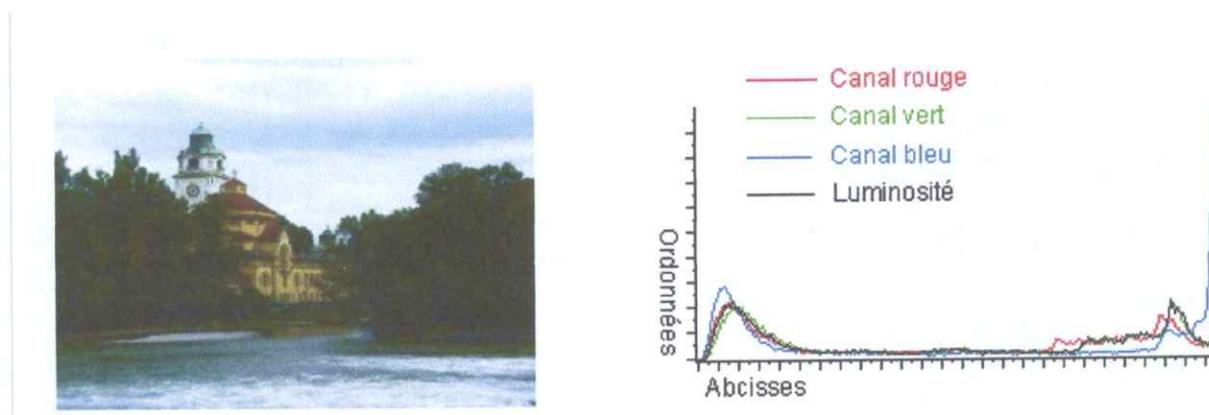


Figure 1.3: histogramme couleur d'une image.

### I.2.2 Contours et régions [4]

- Un contour représente la frontière entre les objets de l'image, ou la limite entre deux pixels dont les niveaux de gris représentent une différence significative.
- Une région est une zone homogène constituée d'un ensemble de pixels adjacents satisfaisants les propriétés de connexité et d'homogénéité.

### I.2.3 Homogénéité [1]

Dans une image, une région est dite homogène si tous les pixels la constituant ont des caractéristiques similaires et uniformes.

Plusieurs facteurs sont prisent en compte pour établir l'homogénéité d'une région, citons par exemple : la couleur, la forme, la texture... etc.

### I.3 FORMATS D'IMAGES NUMÉRIQUES [2,5]

Pour réduire la place occupée en mémoire, on utilise divers algorithmes de compression et donc différents formats de stockage. Il existe plusieurs formats d'images, on distingue parmi eux :

#### **Format BMP (BitMaP)**

C'est un format non compressé, il a été créé par Microsoft et IBM. On peut choisir un codage de 1 à 24 bits par pixel. Ce format prend en charge les modes RVB, Couleurs indexées, niveaux de gris. Il a été conçu pour les ordinateurs personnels (PC) et pour une utilisation dans un environnement Windows et OS/2. Les fichiers BMP ont l'extension .bmp

#### **Format JPEG**

Ce format est développé par *Joint Photographic Expert Group*. C'est un format de compression très efficace mais avec perte de qualité. Cependant, il est mieux adapté aux images de couleurs vraies grâce à sa technique de compression JPEG.

Ce format prend en charge les modes colorimétriques CMJN, RVB et Niveaux de gris. Il est approprié pour les photographies (notamment pour une publication sur Internet) et les images de grandes tailles. Les extensions les plus connues sont : .jpeg, .jpg,

#### **Format GIF (Graphics Interchange Format)**

C'est un format très largement diffusé et répandu. Il a l'extension .gif, il est basé sur l'algorithme LZW, dont les performances ne sont pas excellentes. Il est mieux adapté aux images ne nécessitant pas une grande palette de couleurs (niveaux de gris ou 256 couleurs). Une des applications où son efficacité est prouvée est le Web.

#### **Format TIFF (Tag Image File Format)**

Ce format supporte de monochrome à 16 millions de couleurs, supporte aussi la palette de couleurs. Ainsi, il définit des balises pour décrire les caractéristiques de l'image.

Le format TIFF a les extensions .tif et .tiff, il prend en charge les fichiers CMJN, RVB, à couleurs indexées et en niveaux de gris.

## I.4 LA COULEUR

### I.4.1 Perception de la couleur [6]

La couleur n'est pas une réalité physique, c'est l'impression produite sur notre cerveau, via notre œil, par les radiations lumineuses issues d'un objet.

C'est Newton qui, en 1676, a montré que la lumière blanche du soleil se décompose selon les couleurs du spectre au passage d'un prisme. Chaque composante de la lumière a un coefficient de réfraction différent en fonction de sa longueur d'onde. La lumière visible présente un spectre continu dont les longueurs d'onde varient de 380nm à 780nm.

La surface d'un objet absorbe une partie des radiations lumineuses qu'elle reçoit et réfléchit les autres. Notre œil ne reçoit que les radiations réfléchies. Ce sont ces radiations réfléchies qui vont déterminer la couleur de cet objet. On dira alors qu'un objet est «rouge» parce que sa surface réfléchit la composante rouge de la lumière et absorbe toutes les autres.

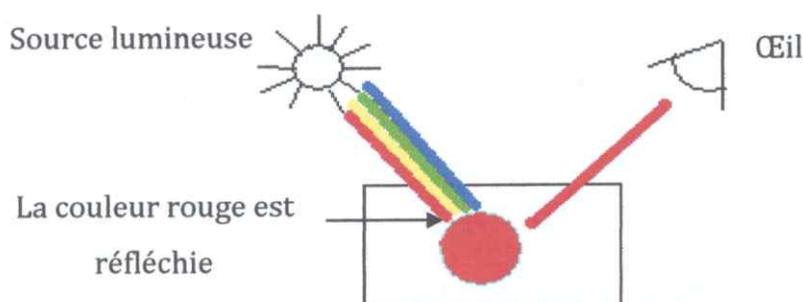


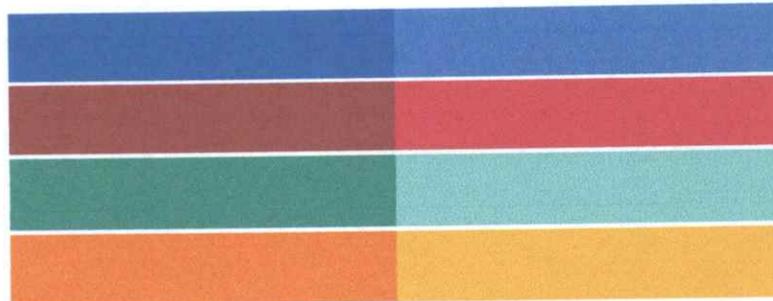
Figure 1.4: perception de la couleur.

### I.4.2 Caractéristiques de la couleur [7,8]

Toute couleur répond toujours aux caractéristiques suivantes : la luminosité, la teinte, la saturation et l'intensité.

### I.4.2.1 Luminosité [8]

Des couleurs peuvent avoir la même tonalité mais paraître plus ou moins claires. Les comparaisons des couleurs qui suivent affichent pour chacune d'entre elles, à gauche une couleur et à droite une variante plus claire de cette couleur:



Pour les couleurs du côté droit 25% de luminosité en plus par rapport au côté gauche ont été fixés. Par la luminosité, les couleurs ne deviennent pas plus vives! Les couleurs soutenues restent soutenues et les couleurs pastel restent pastel.

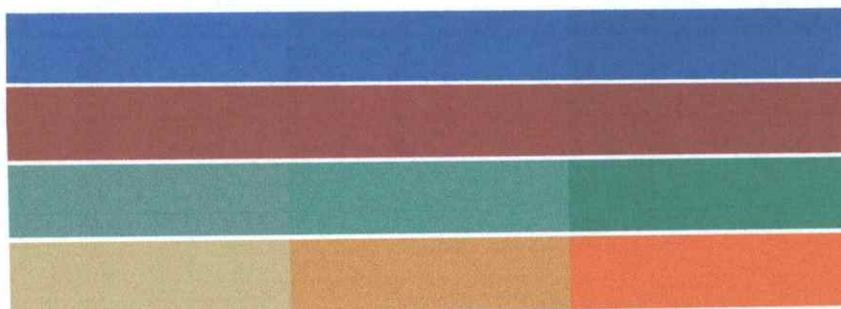
### I.4.2.2 Teinte [8]

Selon CIE (Commission Internationale d'Eclairage), la teinte est définie par la capacité de la vision humaine d'associer une zone de couleur à une autre définie comme étant de couleur rouge, verte ou bleue ou d'une combinaison de deux de ces couleurs.

### I.4.2.3 Saturation [7]

A l'inverse de la luminosité, la saturation influence la perception subjective du caractère vif d'une couleur.

Les comparaisons des couleurs suivantes montrent respectivement au milieu la couleur de base, à gauche une variante avec une saturation moindre et à droite une variante avec une saturation plus élevée.



En ce qui concerne la luminosité, les différents coloris ont été réglés sur la même luminosité. Les couleurs avec une saturation plus ou moins élevée semblent certes souvent au premier coup d'œil plus claires ou plus sombres mais c'est à l'évidence une conclusion fallacieuse.

#### 1.4.2.4 Intensité

Intensité de lumière que la couleur réfléchit/transmet, elle représente la quantité de lumière réfléchi par un objet. C'est ce qui nous permet de dire qu'une couleur est très foncée ou claire.

### 1.4.3 Codage de la couleur [3]

Dans une image, chaque pixel contient une valeur, cette valeur est codée sur un certain nombre de bits déterminant la couleur ou l'intensité du pixel, on l'appelle la profondeur de codage ou la profondeur de couleur, il existe plusieurs standards pour le codage de la profondeur :

- **Bitmap noir et blanc** : chaque pixel est codé sur un seul bit, donc il est possible de définir deux couleurs (noir ou blanc).
  
- **Bitmap 16 couleurs (16 niveaux de gris)** : chaque pixel est codé sur 4 bits, donc 16 couleurs différentes ou bien 16 niveaux de gris allant de noir au blanc.
  
- **Bitmap 256 couleurs (256 niveaux de gris)** : chaque pixel est codé sur 8 bits, donc 256 couleurs (niveaux de gris) allant de noir au blanc.
  
- **Palette de couleurs** : une palette (table de couleurs) contient l'ensemble de couleurs pouvant être contenus dans l'image, à chaque couleur est associé un indice, le nombre de bits réservés pour le codage de chaque indice détermine le nombre de couleurs pouvant être utilisés. Ainsi, en codant les indices sur 8 bits, il est possible de définir 256 couleurs utilisables, c'est-à-dire que chaque case du tableau à deux dimensions représentant l'image va contenir un nombre indiquant l'indice de la couleur à utiliser.

- **Couleurs vrais (True Color)** : dans cette représentation, chaque pixel est représenté par un entier comportant les 3 composantes Rouge, Vert, bleu, chacune codée sur 8 bits, c'est-à-dire 24 bits en total (16 millions de couleurs). On peut ajouter une 4<sup>ème</sup> composante représentant la transparence ou la texture, alors chaque pixel est codé sur 32 bits.

## I.5 ESPACES COULEURS

Pour le codage et la représentation de couleur, l'espace RGB est le plus répandu, mais il a des inconvénients. D'autres espaces sont utilisés pour dissimuler ces inconvénients, on peut citer : l'espace Lab, l'espace CMJN, XYZ, espace TSL, HSV...etc [8] .

On s'intéresse dans notre mémoire aux espaces couleurs RGB, HSV, YCbCr, LUV et l'espace HMMD qu'on va décrire dans le troisième chapitre.

## CONCLUSION

Dans ce chapitre, nous avons vu les notions de base sur l'image et la couleur. Nous avons décrit l'image numérique que nous utiliserons dans l'élaboration de notre système d'indexation et de recherche d'image par le contenu couleur.

Dans le chapitre suivant, nous nous intéressons au principe de l'indexation et de la recherche d'image par contenu.

## Chapitre 2

# *Indexation et Recherche d'images*

## II.2 APPROCHES DE L'INDEXATION

### II.2.1 Approche globale

Cette approche considère l'image dans son ensemble et la caractérise en utilisant des statistiques calculées sur l'image entière. Les différentes approches globales portent généralement sur trois critères : la couleur, la texture et la forme [10].

Pour chaque critère, un grand nombre de descripteurs ont été développés. Ils peuvent ensuite être combinés pour obtenir un descripteur complet et robuste.

- **La couleur [13]**

La couleur est très souvent le premier descripteur qui est employé pour la recherche d'images. Plusieurs travaux ont déjà prouvé qu'il s'agit d'un descripteur efficace (Smeulders, 2000).

Il existe de nombreuses possibilités d'attributs pour caractériser la couleur tels que l'histogramme, les moments couleur... Et on trouve qu'avec la même méthode, si on change l'espace de couleur, il peut émerger des informations différentes de l'image. Cette section se consacre aux méthodes de couleur les plus utilisées dans un système d'indexation et recherche d'images ainsi que sur les espaces de couleur.

- **La texture [10]**

Une texture peut être caractérisée par la répétition d'un motif ou de quelques éléments. Dans le cadre de la recherche par le contenu, elle permet de distinguer des zones de couleurs similaires, mais de sémantique différente (par exemple, le bleu du ciel et le bleu de la mer).

Les caractéristiques de base de la texture les plus utilisées et qui correspondent à la perception visuelle humaine ont été définies par Amadasum, al et Rolland. Il s'agit du contraste, de la complexité, de la grossièreté, de la forme, de la direction et de la force.

## II.4 ARCHITECTURE GENERALE DU SYSTEME D'INDEXATION ET DE RECHERCHE D'IMAGES

Un système d'indexation et de recherche d'images par le contenu est un système qui permet la recherche dans une base d'images en se basant sur les trois attributs couleur, la texture et la forme.

La performance du système de recherche d'images dépend notamment de l'indexation des images qui doit permettre de retrouver la sémantique associée à l'image, du modèle de représentation qui doit être efficace et de la mesure de similarité qui doit permettre de retrouver les documents pertinents.

La figure suivante présente l'architecture du système de recherche d'images par le contenu[12].

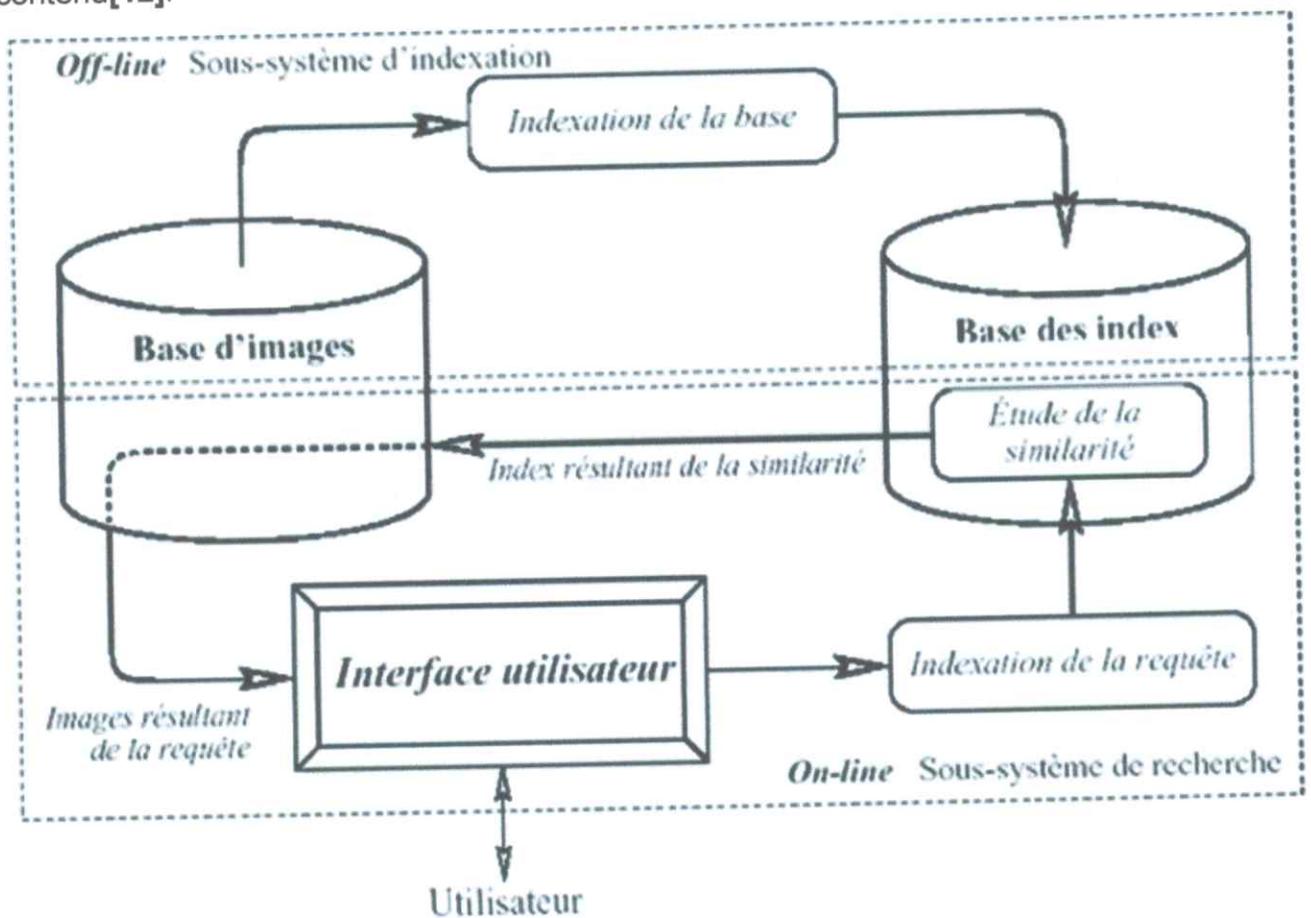


Figure 2.1 : L'architecture d'un système d'indexation et recherche d'images par le contenu.

## CONCLUSION

La recherche d'images par le contenu est devenue utile. Donc, le besoin de systèmes capables de décrire les différents types d'informations dans les images se fait de plus en plus fortement sentir.

Plusieurs standards utilisés pour décrire le contenu d'images, MPEG-7 est l'un de ces standards.

# Chapitre 3

*Les descripteurs couleurs*

*MPEG-7*

## INTRODUCTION

Il existe divers descripteurs d'images : la couleur, la texture, la forme etc... Ces descripteurs peuvent être extraits par diverses approches en traitant le contenu de l'image (segmentation, compression...). De ces approches, sont nés des standards de description d'images.

Nous avons choisis le MPEG-7 car il est un standard ISO/IEC [15], et les bases multimédias existantes sur Internet seront décrites en MPEG-7 dans un futur proche. Dans ce chapitre, nous présentons le standard MPEG-7, en particulier les différents descripteurs couleurs proposés par ce standard. Ce sont les descripteurs que l'on va considérer par la suite dans notre travail et pour nos besoins expérimentaux.

### III.1 DÉFINITION DE MPEG-7 [15]

Le MPEG-7, formellement appelé « *Multimédia Content Description Interface* » est un standard ISO/IEC développé en automne 2001 par le groupe MPEG.

Le MPEG-7 permet de décrire le contenu d'un fichier multimédia avec un certain niveau d'interprétation des informations de ce fichier. On peut accéder à ces informations grâce à un équipement spécifique ou un logiciel, depuis un ordinateur.

Le MPEG7 n'est donc pas dédié à une architecture, à un média particulier, mais permet de standardiser un nouveau moyen de recherche multimédia et ce pour un très grand nombre d'applications.

A la différence de MPEG-1, MPEG-2 et MPEG-4, qui standardisent l'encodage même des documents audiovisuels, MPEG-7 régit la description de ceux-ci et non pas l'encodage. La navigation, la recherche et le filtrage peuvent donc s'opérer sur ces descriptions plutôt que sur les fichiers sources, ce qui évite de devoir décompresser ces derniers et de les traiter à chaque requête.

### III.2 FONCTIONNALITÉS MAJEURES DU MPEG-7 [15]

Les informations qui apparaîtront dans un document MPEG-7 seront de 5 natures différentes, résumées dans le tableau suivant :

Ensemble des éléments	Fonctionnalité
<b>Création et production</b>	Des méta-informations qui décrivent la création et la production du contenu, elles décrivent le titre, le créateur, le but de la création.
<b>Utilisation</b>	Des méta-informations reliées à l'utilisation du contenu: Elles comportent les droits d'accès, des informations financières, des droits de publication. Ces informations peuvent faire l'objet de changement durant la durée de vie du contenu audio-visuel.
<b>Média</b>	Ces informations décrivent les caractéristiques de stockage: Format, éléments pour identifier le média.
<b>Aspects structurels</b>	Des descriptions d'un point de vue contenu: Ces informations décrivent les segments qui peuvent représenter des composantes spatiales, temporelles ou spatio-temporelles du contenu audio-visuel. Chaque segment peut être décrit par les caractéristiques suivantes (la couleur, la texture, la forme, la motion, d'autres caractéristiques audio...) et quelques informations sémantiques élémentaires.
<b>Aspect Conceptuels</b>	Des descriptions du contenu audio-visuel d'un point de vue conceptuel. Ces informations ne sont pas indiquées dans les documents techniques du standard MPEG, car elles sont en cours de standardisation.

Tableau 3.1 : fonctionnalités majeures du standard MPEG-7

### III.3 PORTÉE DE STANDARD MPEG-7[18]

Fidèle à sa politique, la norme MPEG définit ses standards dans le but d'atteindre une interopérabilité maximale entre les produits des différents acteurs du marché, tout en permettant la plus grande concurrence entre ces acteurs.

La figure suivante schématise la chaîne de traitements impliqués dans les processus de recherche et de sélection des contenus multimédia.

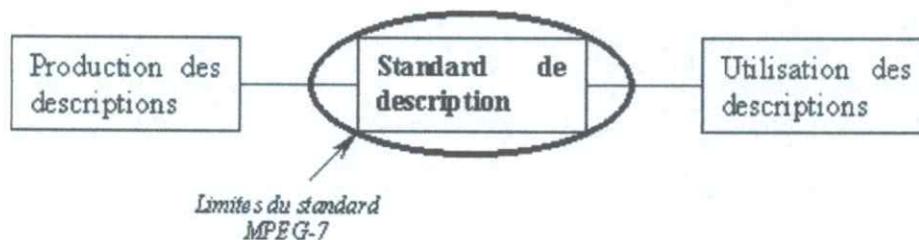


Figure 3.1 : place du standard MPEG-7 dans le processus d'accès au contenu.

### III.4 LES APPLICATIONS MPEG-7

Différentes applications ont déjà été identifiées par les acteurs de MPEG-7 parmi lesquelles on distingue [18]:

- ✓ Les applications de type "Pull" : la consultation des bases vidéos.
- ✓ Les applications de type "Push" : le filtrage dans des bouquets numériques diffusés (TV par satellite).
- ✓ Les applications spécialisées : le téléachat, l'édition multimédia, etc.

### III.5 STRUCTURES DE MPEG-7

MPEG-7 propose plusieurs structures d'information permettant d'exprimer les descriptions. Ces structures sont les suivantes [15, 16]:

- **Les descripteurs (D)** : permettent de décrire les différentes caractéristiques des contenus, notamment visuels et sonores. Chaque descripteur définit une caractéristique élémentaire, selon une syntaxe et une sémantique précise et propre à chaque caractéristique.

- **Les schémas de description (SD)** : En règle générale, plusieurs descripteurs sont nécessaires à la description d'une ressource. Les schémas de description (*Description Schemes* en Anglais) permettent de regrouper ces descripteurs selon une structure et une sémantique bien définie. Les SD peuvent ensuite être regroupés entre eux et avec d'autres descripteurs afin de produire des schémas de description de niveau supérieur.
- **Langage de définition des descriptions (LDD)**: basé sur XML, il permet de créer de nouveaux schémas de description et, peut être, de nouveaux descripteurs, il doit permettre aussi la modification et l'extension modifier des descripteurs et des schémas de description existants. Il utilise les types primitifs définis par XSchema.
- **Valeur d'un descripteur** : Instance d'un descripteur pour un ensemble de données.
- **Outils de description** : le regroupement des Descripteurs et les schémas de description créé un ensemble d'outils de description (*Description Tools* ou *Multimedia Description Schemes* en Anglais). Ces outils permettent de générer les descripteurs et les schémas de descriptions du standard MPEG-7, de les gérer, les manipuler...etc.

Le schéma suivant représente la hiérarchie de MPEG7, c'est-à-dire la relation entre les descripteurs ( $D_s$ ), les schémas de description ( $SD_s$ ) et le langage de définition des descriptions (LDD).

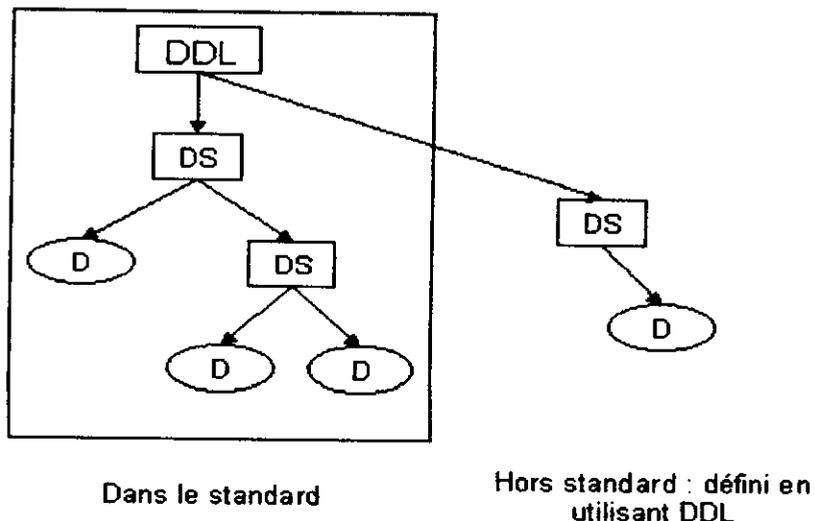


Figure 3.2 : diagramme de hiérarchie de MPEG7.

### III.6 LES DESCRIPTEURS VISUELS MPEG-7 [16]

Tels que spécifiés dans la norme MPEG-7, les descripteurs visuels peuvent se classer en quatre catégories distinctes : les descripteurs couleurs, les descripteurs de texture, les descripteurs de forme et les descripteurs de mouvement.

- Les **descripteurs couleurs** (*Color Descriptors*) servent de caractériser les images en fonction des différentes couleurs qu'elles contiennent. Il existe six descripteurs qu'on va détailler ci-dessous.
- Les **descripteurs de texture** (*Texture Descriptors*) permettent d'analyser les différents motifs ou textures présents sur une image. Chaque élément d'une image a en effet une texture qui lui est propre  
Le standard MPEG-7 propose trois descripteurs de texture : *Homogeneous Texture*, *Edge Histogram* et *texture Browsing Descriptor*.
- Les **descripteurs de forme** (*Shape Descriptors*) permettent de reconnaître les objets d'une image en se basant sur leurs formes plus précisément sur les régions et les contours des objets.

MPEG-7 propose les descripteurs de formes suivants : *region-based descriptor*, *contour-based descriptor*, *2D/3D shape descriptor*, *3D shape descriptor*.

- Enfin, les **descripteurs de mouvement** (*Motion Descriptors*) permettent d'ajouter une composante plus dynamique aux descripteurs précédents, en exprimant les informations de mouvement des objets entre plusieurs plans lors de la description de segments vidéo.

Les différents descripteurs de mouvement sont : *Motion Activity Descriptors*, *Camera Motion Descriptors*, *Motion Trajectory Descriptors* et *Parametric Motion Descriptors*.

### III.7 LES DESCRIPTEURS COULEURS MPEG-7

La couleur est un attribut visuel très important pour la vision humaine et aux traitements réalisés par l'ordinateur.

Différents facteurs sont prisent en compte pour la sélection des différents descripteurs couleurs, on distingue :

- L'aptitude à caractériser au mieux la similarité perceptuelle de la couleur.
- Basse complexité des techniques de calcul et de comparaison des descripteurs.
- La taille de description qui joue un rôle très important dans l'indexation et la transmission des descripteurs hors les réseaux à largeur de bande limitée.
- Le passage à l'échelle et l'interopérabilité des descripteurs [17].

Le standard MPEG-7 définit généralement six descripteurs couleurs, le descripteur *Color Space* qui définit les différents espaces couleurs utilisés par les autres descripteurs, les trois descripteurs *Scalable Color*, *Color Structure* et *Group Of Pictures or Group OF frames* qui sont dérivés d'un histogramme couleur, ainsi que le descripteur *Dominant Color* et le descripteur *Color Layout* [19].

Les différents descripteurs, avec leurs espaces couleurs, sont affichés dans le schéma suivant [19]:

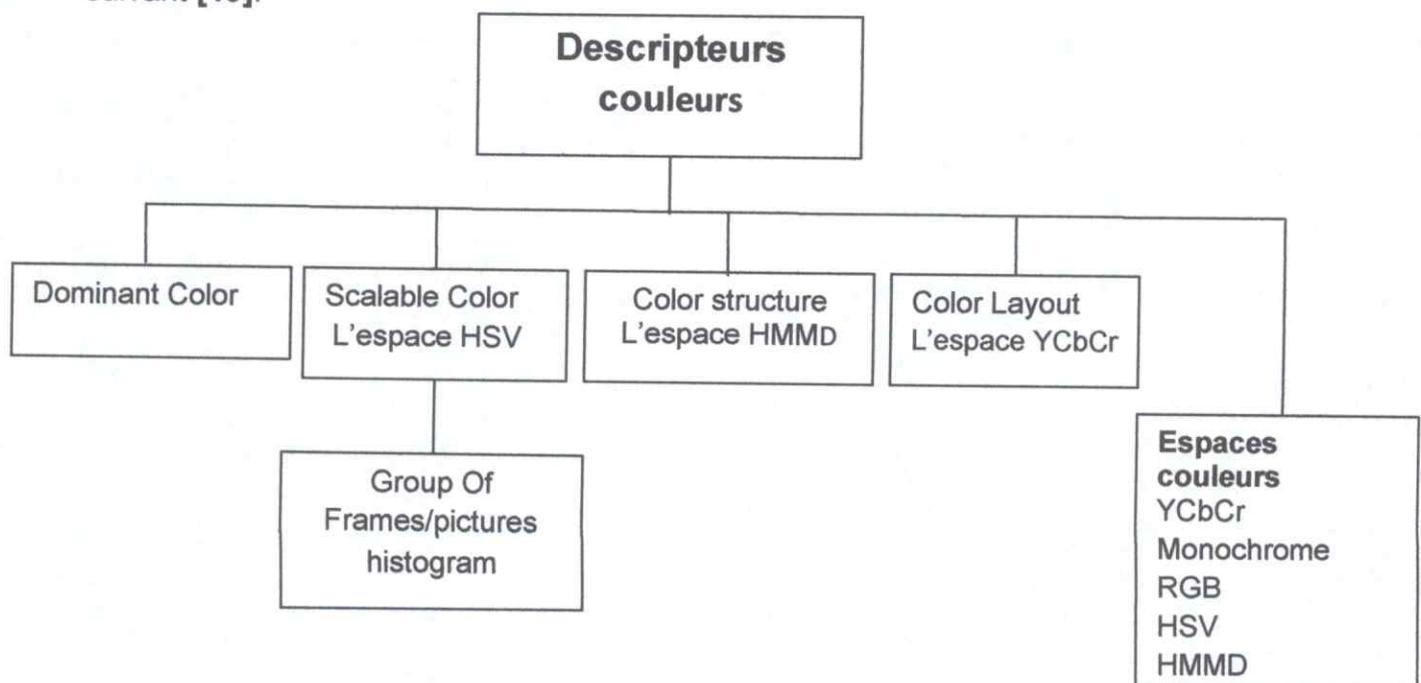


Figure 3.3 : les différents descripteurs couleurs.

### III.7.1 Le descripteur Color Space

Le descripteur espaces couleurs (*Color space*) spécifie un ensemble d'espaces couleurs pour être utilisés dans un autre descripteur couleur. Les différents espaces couleurs utilisés par le standard MPEG-7 sont l'espace monochrome, l'espace RGB, HSV,  $YCbCr$  et le nouveau espace HMMD. L'espace LUV, ainsi que l'espace monochrome (intensité seulement), qui correspond à la composante Y de  $YCbCr$  sont aussi supportés [19].

Le descripteur Color Space associe un autre descripteur qui est le descripteur de quantification des couleurs (*Color Quantization Descriptor*) qui permet de spécifier le partitionnement de l'espace couleur donné en valeurs discrètes (bins) [17].

#### III.7.1.1 L'espace couleur RGB (Red, Green, Blue) [16, 8]

C'est l'espace couleur le plus couramment utilisé, dans lequel toute couleur est exprimée en fonction de ses composantes rouge (R), verte (G) et bleue (B). Une valeur entre 0 et 255 indique pour chaque composante l'intensité de la couleur sélectionnée, 0 signifiant que la couleur est absente et 255 indiquant que la couleur est présente à 100%.

Généralement, on présente l'espace RGB par un cube dont chacun des axes correspond à une couleur primaire (figure 4.a).

le mode RGB est utilisé dans tous les moniteurs, utilisé généralement pour l'affichage. C'est un mode de couleur additif (figure 4.b), Par exemple :

Rouge+vert = jaune  
 Rouge+vert+bleu = blanc  
 Vert+bleu = cyan

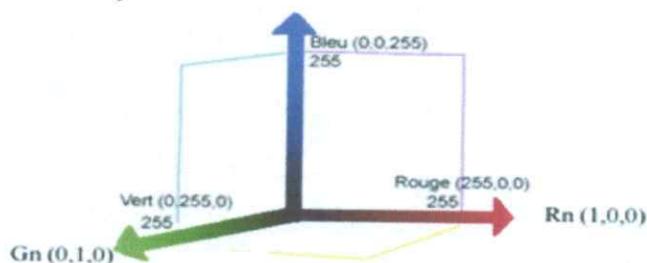


Figure 4.a

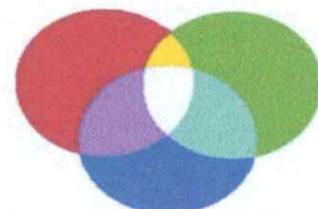


figure 4.b

Figure 3.4 : 3.4.a) représentation cube de l'espace couleur RGB.

3.4.b) Mode additif de RGB.

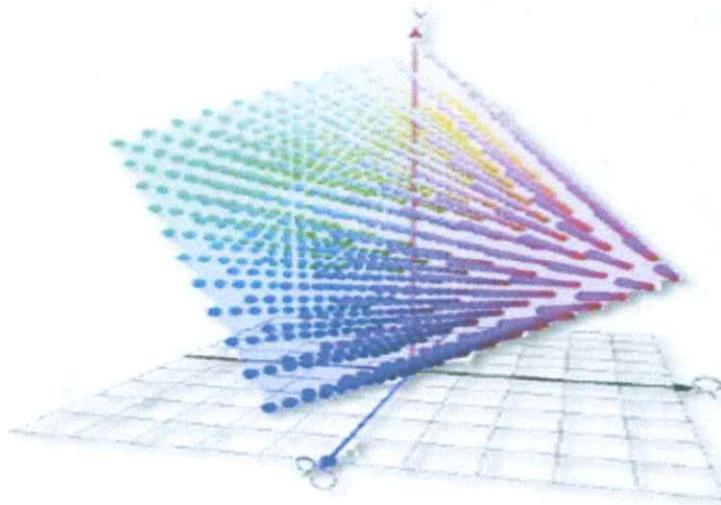
### III.7.1.2 L'espace couleur $YCbCr$ [16, 17]

C'est l'ancien espace couleur utilisé par les précédents standards MPEG, ces trois dimensions sont la luminance ( $Y$ ), la chrominance rouge ( $C_r$ ) et la chrominance bleue ( $C_b$ ).

Il est défini par une transformation linéaire de RGB comme suit :

$$\begin{aligned} Y &= 0.299 * R + 0.587 * G + 0.114 * B \\ C_b &= -0.169 * R - 0.331 * G + 0.500 * B \\ C_r &= 0.500 * R - 0.419 * G - 0.081 * B \end{aligned}$$

La répartition des couleurs sur les trois axes  $Y$ ,  $C_b$  et  $C_r$  est montrée dans la **figure 3.5**, le paramètre  $Y$  définit la luminosité alors que les deux autres paramètres gèrent la coloration.



**Figure 3.5** : représentation des couleurs dans l'espace  $YCbCr$ .

### III.7.1.3 L'espace couleur HSV (Hue, Saturation, Value) [16,17]

Cet espace est défini comme un cylindre qui se compose d'une teinte ( $H$ ), d'une saturation ( $S$ ) et d'une valeur ( $V$ ).

La composante teinte permet de choisir la teinte générale de la couleur, elle est représentée par un angle de  $0$  à  $360^\circ$ . La saturation permet ensuite de varier la pureté ou l'intensité de cette teinte, entre  $0$  et  $1$ . Enfin, le paramètre valeur (entre  $0$  et  $1$ ) définit la brillance de la couleur et permet de rendre une couleur plus sombre ou plus claire.

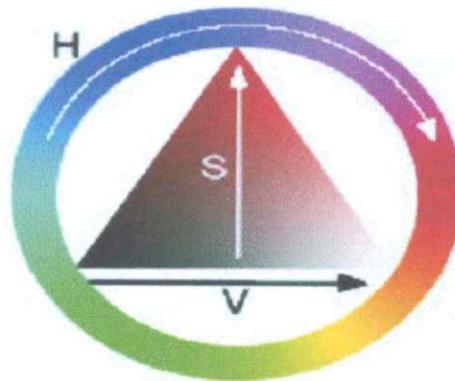


Figure 3.6 : représentation des couleurs dans l'espace HSV.

Les trois composantes sont exprimées par une transformation non linéaire des trois composantes de RGB comme suit :

```

Max = max (R, G, B);
Min = min (R, G, B);
Value = Max;
if (Max == 0) then Saturation = 0;
else
Saturation = (Max - Min)/Max;
if (Max == Min) Hue = 0 ; // couleur
achromatique.
otherwise:
if (max == R && G >= B)
Hue = 60*(G - B)/(Max - Min)
else if (Max == R && G < B)
Hue = 360 + 60*(G - B)/(Max - Min)
else if (G == Max)
Hue = 60*(2.0 + (B - R)/( Max - Min))
else
Hue = 60*(4.0 + (R - G)/( Max - Min))

```

#### III.7.1.4 L'espace couleur LUV [16, 17]

L'espace couleur LUV vise à une représentation uniforme des couleurs. L'idée principale de cet espace est de rapprocher autant que possible l'expression des couleurs à l'écran de leur perception par l'œil humain.

La Figure 3.7 présente une illustration de cet espace de couleurs. On remarque sur les bords gauche et supérieur que toutes les couleurs du spectre lumineux sont définies.

Le bord droit contient quand à lui les couleurs extra spectrales, telles le magenta ou le violet.

Cet espace est en particulier utilisé dans les écrans de TV et d'ordinateurs, qui combinent en général des sources rouges, vertes et bleues, c'est-à-dire les trois couleurs de base de l'espace RGB pour créer toutes les couleurs à afficher.

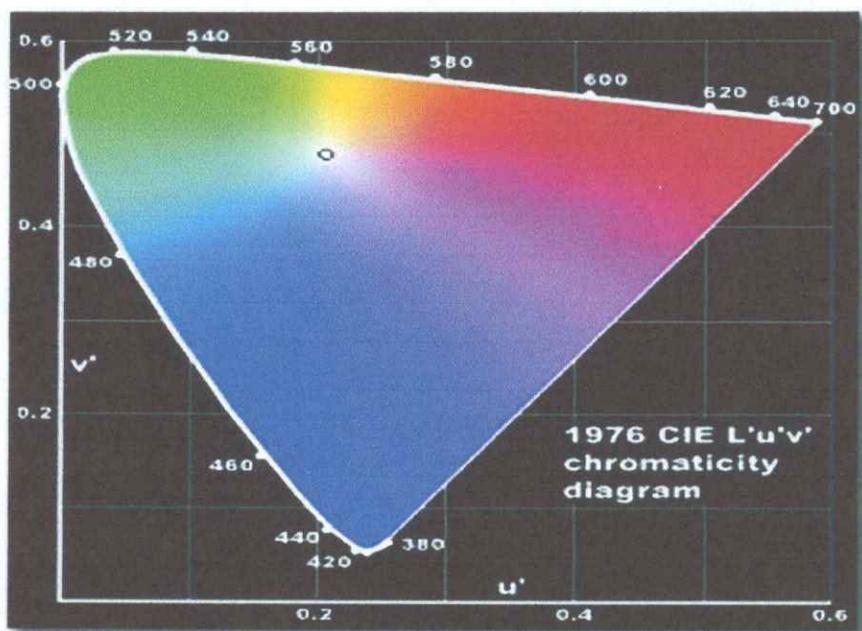


Figure 3.7: l'espace de couleurs LUV.

### III.7.1.5 L'espace couleur HMMD (Hue, Max, Min, Diff) [17]

C'est un espace couleur perceptuel et uniforme, la forme double cône confine cet espace comme il est affiché dans la **figure 3.8**. Tel que les composantes Max, Min et Diff sont obtenues par les équations de transformation entre RGB et HMMD comme suit :

$$\text{Max} = \max (R, G, B).$$

$$\text{Min} = \min (R, G, B).$$

$$\text{Diff} = \text{Max} - \text{Min}.$$

Une 5<sup>ème</sup> composante Sum peut être définie comme suit :  $\text{Sum} = (\text{Max} + \text{Min}) / 2$ .

Donc, l'espace HMMD est généralement décrit par 5 composantes. Cependant, un ensemble de trois composantes, soit {Hue, Max, Min} ou {Hue, Diff, Sum} est suffisant pour former cet espace, et spécifier un point couleur.

Dans l'espace HMMD, la composante Hue (varie entre  $0^\circ$  et  $360^\circ$ ) à la même propriété que Hue dans l'espace HSV. Max (entre 0 et 1) indique combien la couleur noire est présente, donnant la saveur de la nuance ou de l'obscurité (*blackness* en anglais). Min (entre 0 et 1) indique combien la couleur blanche est présente, donnant la saveur de la pureté (*whiteness* en anglais). Diff (entre 0 et 1) indique combien une couleur est proche des couleurs pures, donnant la saveur de ton ou de coloration (*colorfulness* en anglais). Enfin, la composante Sum (ente 0 et 1) indique la clarté (*brightness* en anglais) de la couleur.

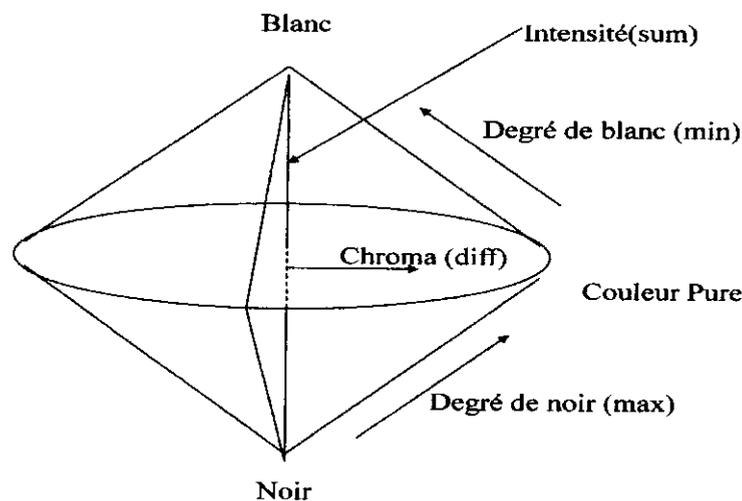


Figure 3.8 : représentation double cônes de l'espace couleur HMMD.

### III.7.2 Le descripteur Dominant Color [16, 17, 19]

- Définition

Le descripteur de couleurs dominantes (*Dominant Color*), noté DCD représente de manière compacte les différentes couleurs présentes dans une image, ou une région d'une image. Il permet la recherche par similarité dans des bases de données d'images ainsi que le parcours de ces bases de données à partir d'une ou plusieurs valeurs couleurs.

L'utilisation de DCD sur des images entières permet de retrouver d'autres images contenant les mêmes couleurs, indépendamment de leur emplacement dans l'image.

Le descripteur *Dominant Color* est défini comme suit :

$$F = \{(c_i, p_i, v_i), s\}, \quad (i = 1, 2, \dots, N)$$

Avec :

$N$  représente le nombre de couleurs dominantes (8 couleurs dominantes au maximum).

$c_i$  représente la  $i^{\text{ème}}$  couleur dominante (c'est un vecteur qui représente les valeurs des composantes de l'espace couleur 3D dans RGB).

$p_i$  est le pourcentage de la  $i^{\text{ème}}$  couleur dominante (c'est la fraction des pixels de l'image correspondant à  $c_i$ ).

$v_i$  est la variance couleur de  $c_i$  qui décrit la variation des valeurs couleurs de pixels dans un cluster autour de la couleur dominante correspondante.

La valeur unique  $s$  représente la cohérence spatiale (nombre qui représente l'homogénéité spatiale des couleurs dominantes dans l'image).

- **Extraction de DCD**

Le processus d'extraction de DCD est effectué par un algorithme de clustering, qui est l'algorithme de Lloyd généralisé. Cette extraction se fait dans un espace de couleurs à perception uniforme, tel que l'espace LUV.

La fonction distorsion dans le  $i^{\text{ème}}$  cluster est donnée par :

$$D_i = \sum_n h(n) \|x(n) - c_i\|^2, \quad x(n) \in C_i \quad (1)$$

Tel que :

$c_i$  est le centre de masse de cluster  $C_i$ .

$x(n)$  est le vecteur couleur au pixel  $n$ .

$h(n)$  définit le poids perceptuel lié au pixel  $n$ .

La règle de mise à jour pour la métrique de distorsion ci-dessus peut être dérivée pour être :

$$c_i = \frac{\sum h(n)x(n)}{\sum h(n)}, \quad x(n) \in C_i \quad (2)$$

Au début, la procédure est initialisée avec un seul cluster comprenant l'ensemble des pixels de l'image et une couleur représentative de ce cluster est calculée comme un centre de masse.

A chaque itération, l'algorithme recalcule les clusters de couleurs et calcule une fonction de distorsion. Ceci se fait jusqu'à ce qu'un critère d'arrêt est atteint (distorsion minimale ou nombre maximal d'itérations).

Des perturbations sont ensuite ajoutées aux centres de masse des clusters ayant les distorsions les plus grandes afin de les faire diminuer. L'algorithme s'arrête alors lorsque la distorsion maximale est inférieure à un certain seuil, ou que le nombre maximal de clusters est atteint.

Le pourcentage  $p_i$  de pixels de l'image appartenant à chacune des couleurs dominantes est ensuite calculé, le résultat est uniformément quantifié sur 5 bits. Les valeurs couleurs sont ensuite quantifiées sur 1 à 12 bits en fonction de l'espace couleur et les descripteurs *Color Quantization* associés.

Le DCD détermine ainsi la cohérence spatiale des clusters, l'algorithme d'extraction détermine parmi les voisins directs de chaque pixel, combien appartiennent au même cluster. Un masque de 3x3 pixels appliqués successivement sur chaque pixel est utilisé pour ce but. La moyenne des valeurs pour chaque pixel d'un cluster donne la cohérence spatiale liée à la couleur dominante qu'il représente.

La cohérence spatiale générale de l'image est ainsi calculée en combinant linéairement les cohérences spatiales des différents clusters avec les pourcentages  $P_i$  qui correspondent et qui représentent les poids utilisés pour la combinaison. Cette cohérence spatiale est ainsi quantifiée non linéairement sur 5 bits. Les valeurs de 1 à 31 indiquent l'intensité de la cohérence, dans l'ordre croissant tandis que la valeur 0 indique que la valeur n'a pas été calculée.

- **Représentation**

La syntaxe binaire de descripteur DCD spécifie 3 bits pour représenter le nombre de couleurs dominantes dans l'image (chacune de ces valeurs est identifiée par sa valeur dans l'index des couleurs, en utilisant entre 1 et 12 bits) et le pourcentage de pixels qui lui est associé est codé sur 5 bits. Les variances couleurs optionnelles sont codées sur 1 bit par couleur, la quantification de cette couleur se fait de manière non uniforme.

Le tableau suivant représente la syntaxe binaire ou bien la représentation de descripteur DCD :

Champ	Nombre de bits	Description
NumberOfColors	3	Nombre de couleurs dominantes
SpatialCoherency	5	Cohérence spatiale des couleurs
Percentage []	5	Pourcentage normalisé associé à chaque couleur dominante
ColorVariance [] []	1	Variance pour chaque couleur dominante
Index [] []	1-12	Valeurs des couleurs dominantes

Tableau 3.2 : Représentation du descripteur Dominant Color.

Le descripteur *Dominant Color* peut également représenté dans un document XML comme suit :

```

xsi:type="DominantColorType"><SpatialCoherency>23</SpatialCoherency><Value>
<Percentage>1</Percentage><Index>45 38 44</Index><ColorVariance>1 0
0</ColorVariance></Value><Value><Percentage>10</Percentage><Index>191 186
163</Index><ColorVariance>0 0
0</ColorVariance></Value><Value><Percentage>9</Percentage><Index>229 230
218</Index><ColorVariance>0 0
0</ColorVariance></Value><Value><Percentage>3</Percentage><Index>111 108
111</Index><ColorVariance>0 0
0</ColorVariance></Value><Value><Percentage>4</Percentage><Index>158 154
147</Index><ColorVariance>0 0
0</ColorVariance></Value><Value><Percentage>1</Percentage><Index>117 100
84</Index><ColorVariance>0 0 0</ColorVariance>
</Value></VisualDescriptor>

```

Figure 3.9 : exemple d'un descripteur Dominant Color représenté sous format XML.

- **Correspondance (distances de similarité)**

On considère les deux descripteurs Dominant Color suivants :

$$F_1 = \{(c_{1i}, p_{1i}, v_{1i}), s_1\}, \quad (i = 1, 2, \dots, N_1) \text{ et}$$

$$F_2 = \{(c_{2i}, p_{2i}, v_{2i}), s_2\}, \quad (i = 1, 2, \dots, N_2).$$

La correspondance entre les deux descripteurs F1 et F2 se fait de la manière suivante :

On commence à définir le coefficient de similarité  $a_{k,l}$  entre deux couleurs  $C_k$  et  $C_l$  comme suit :

$$a_{k,l} = \begin{cases} 1 - dk,l / d_{max} & dk,l \leq T_d \\ 0 & dk,l > T_d \end{cases}$$

Où :

- $d_{k,l} = \|C_k - C_l\|$  est la distance Euclidienne entre les deux couleurs  $C_k$  et  $C_l$ .  $T_d$  est la distance maximale pour laquelle deux couleurs sont considérées comme similaires et  $d_{max} = \alpha T_d$ .
- Une valeur recommandée de  $T_d$  est entre 10 et 20 dans l'espace couleur CIE-LUV et la valeur de  $\alpha$  comprise entre 1 et 1.5.

En ignorant le paramètre variance et la cohérence spatiale, le carré de la dissimilarité entre les deux descripteurs précédents est défini comme suit :

$$D^2(F_1, F_2) = \sum_{i=1}^{N_1} p_{1i}^2 + \sum_{j=1}^{N_2} p_{2j}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{1i,2j} p_{1i} p_{2j} \quad (3)$$

Une variation de la distance précédente est d'utiliser le champ de la cohérence spatiale, par exemple, dans les expériences MPEG-7, la distance suivante est utilisée :

$$D_s = w_1 \text{abs}(s_1 - s_2)D + w_2 D \quad (4)$$

Où  $s_1$  et  $s_2$  sont les cohérences spatiales des deux descripteurs,  $w_1$  et  $w_2$  sont des poids fixes ayant respectivement les valeurs 0.3 et 0.7. Et  $D$  représente la dissimilarité présentée ci-dessus.

Lors de l'extraction du descripteur, des mesures de variance peuvent encore être effectuées. Dans ce cas, le calcul de la distance change encore.

### III.7.3 Le descripteur Color Layout [16,17,19]

- **Définition**

Le descripteur *Color Layout* (CLD) représente de manière très compacte la distribution spatiale des couleurs présentes dans une image. Il constitue une variante du descripteur *Dominant Color*, ou l'on aurait d'abord décomposé l'image avant d'extraire les couleurs dominantes région par région.

Le CLD ne dépend pas de la résolution des images analysées et comparées, ce qui le rend particulièrement adapté pour des applications pour lesquelles les problèmes de stockage ou de bande passante sont importants.

L'avantage de ce descripteur est qu'il permet la recherche basée sur les esquisses, ainsi, plus besoin de fournir une image de référence pour effectuer une recherche, l'utilisateur peut créer une esquisse de ce qu'il recherche, presque de la même manière qu'il rentrerait des mots clés dans un moteur de recherche textuelle.

- **Extraction**

Le processus d'extraction de CLD passe par 4 étapes : le partitionnement de l'image, la détection des couleurs représentatives (dominantes), la transformation des valeurs des couleurs et la quantification des résultats (**figure 3.10**). Ainsi, pour exprimer les différentes couleurs de l'image à analyser, on utilise l'espace couleur  $YC_bC_r$ .

- ✓ **Partitionnement d'image** : l'image d'entrée est décomposée en 8x8 sous-images (blocs) de tailles égales.
- ✓ **Détection des couleurs représentatives** : pour chaque bloc, une unique couleur est sélectionnée en utilisant n'importe quelle méthode, mais il est recommandé de retenir la couleur moyenne.

Ces deux premières étapes ont pour résultat une « icône » de 8x8 couleurs.

- ✓ **Transformation DCT** : les trois composantes (Y, Cb et Cr) sont ensuite analysées indépendamment l'une de l'autre en appliquant une transformée en cosinus discrète (DCT), alors 3 ensembles de 64 coefficients DCT sont obtenus.

- ✓ **Quantification des résultats** : les résultats des transformées sont quantifiés de manière non linéaire en utilisant 64 niveaux pour la composante continue (DC) et 32 niveaux pour les coefficients des fonctions en cosinus (AC).

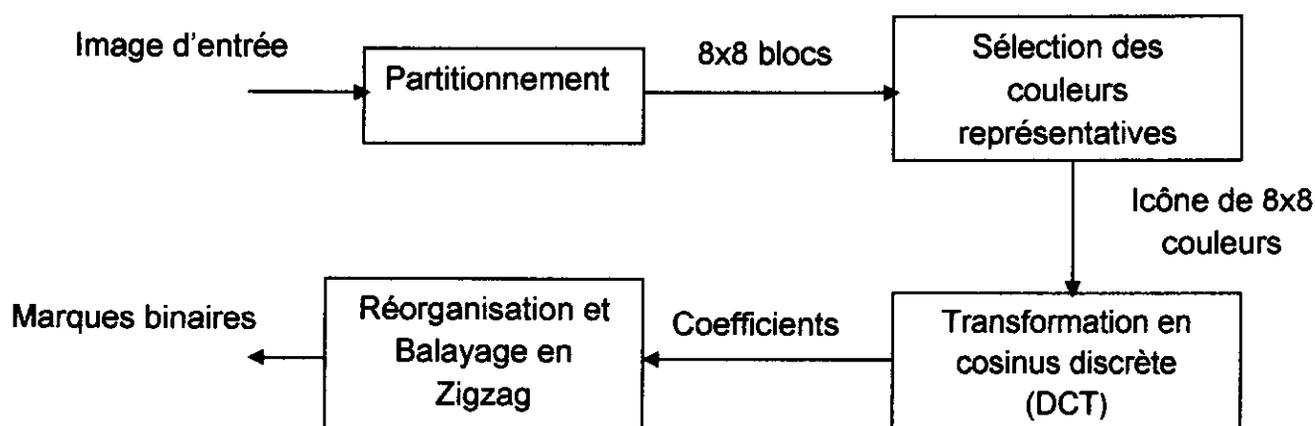


Figure 3.10 : le processus d'extraction de CLD.

#### • Représentation

Le nombre de coefficients utilisés lors du calcul de la DCT est variable, il est représenté dans un champ spécial appelé *CoefficientsPattern field* qui peut avoir 3 valeurs possibles :

- ✓ La première valeur indique l'utilisation de 6 coefficients pour la luminance et 3 pour chaque chrominance
- ✓ La deuxième indique l'utilisation de 6 coefficients pour la luminance et les chrominances.
- ✓ La troisième indique l'utilisation d'un nombre différent de coefficients,

Dans les deux premier cas, les paramètres *NumberofYCoeff* et *NumberofCCoeff* sont fixés alors que dans le troisième, ils prennent une des valeurs suivantes : 3,6,10,15,21,28 ou 64.

Enfin, pour chaque composante, les valeurs des coefficients sont stockées dans un tableau dont la longueur varie en fonction du nombre de coefficients.

Champ	Nombre de bits	Description
CoefficientPattern	1-2	Nombre de coefficients de la TCD.
NumberOfYCoeff	3	Nombre de coefficients de la TCD pour la luminance.
NumberOfCCoeff	3	Nombre de coefficients de la TCD pour les chrominances.
Ycoeff []	5-6	Valeurs des coefficients de la TCD pour la luminance.
CbCoeff []	5-6	Valeurs des coefficients de la TCD pour la chrominance bleue
CrCoeff []	5-6	Valeurs des coefficients de la TCD pour la chrominance rouge.

**Tableau 3.3** : représentation de descripteur Color Layout.

Le descripteur *Color Layout* peut être représenté sous format XML de la manière suivante :

```
<VisualDescriptor xsi:type="ColorLayoutType">
  <YDCCoeff>48</YDCCoeff>
  <CbDCCoeff>25</CbDCCoeff>
  <CrDCCoeff>34</CrDCCoeff>
  <YACCCoeff5>15 26 22 16 14 </YACCCoeff5>
  <CbACCCoeff2>18 17 </CbACCCoeff2>
  <CrACCCoeff2>15 14 </CrACCCoeff2>
</VisualDescriptor>
```

**Figure 3.11** : descripteur Color Layout représenté sous format XML.

- **Correspondance**

Soit les deux descripteurs *Color Layout* CLD {DY, DCb, DCr} et CLD' {DY', DCb', DCr'}, la distance entre ces deux descripteurs est exprimée par l'équation suivante :

$$D = \sqrt{\sum_i w_{yi}(DY_i - DY'_i)^2} + \sqrt{\sum_i w_{bi}(DCb_i - DCb'_i)^2} + \sqrt{\sum_i w_{ri}(DCr_i - DCr'_i)^2} \quad (5)$$

Tel que :

- $DY_i$ ,  $DCb_i$ ,  $DCr_i$  indiquent respectivement les coefficients de la luminance, de la chrominance bleue et de la chrominance rouge.
- L'indice «  $i$  » qui représente l'ordre de balayage en Zigzag pour les coefficients.
- $W_{yi}$ ,  $W_{bi}$ ,  $W_{ri}$  sont des poids associés à chaque coefficient de chaque composante, on associe les plus grands poids pour les composantes qui ont la plus basse fréquence.

La distance euclidienne entre les coefficients est calculée pour chaque composante, en tenant compte les poids associés à chaque coefficient. La somme des distances des trois composantes constitue la distance totale entre deux descripteurs. Si deux descripteurs n'ont pas le même nombre de coefficients, on ne considère que les coefficients communs.

Comme le descripteur CLD peut être appliqué à une région image, il est possible d'avoir des pixels vides. Dans ce cas, il faut d'abord remplir ces pixels vides par la couleur moyenne de l'image avant de calculer ce descripteur et de pouvoir le comparer à d'autres descripteurs.

### III.7.4 Le descripteur Scalable Color [16, 17, 19]

- **Définition**

Le descripteur *Scalable Color* (SCD) peut être vu comme un histogramme des couleurs présentes dans l'image, cet histogramme est ensuite encodé en utilisant la transformée de Haar. Le calcul de l'histogramme des couleurs se fait en utilisant l'espace des couleurs HSV et la quantification des valeurs des différentes composantes se fait linéairement.

- **Extraction**

Pour calculer le SCD, on commence par analyser l'image et extraire un histogramme. Les 256 valeurs (bins) de cet histogramme sont extraites, normalisées et non linéairement représentées par une valeur entière sur 4 bits, cette représentation est faite d'une manière non linéaire afin de donner plus de signification aux valeurs plus faibles. La transformée de Haar est appliquée pour ces valeurs de 4 bits à travers les bins d'histogramme.

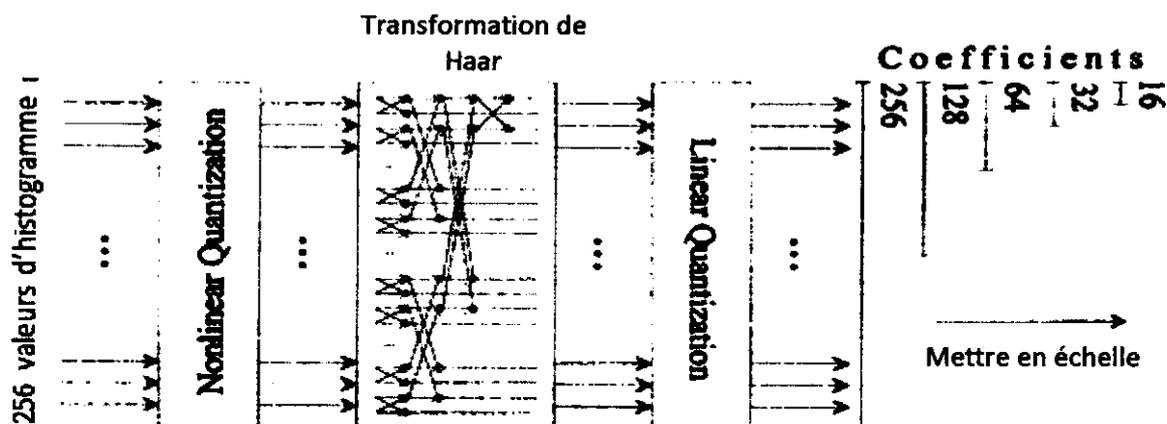


Figure 3.12 : Génération du descripteur Scalable Color.

Lors de chaque passage, les coefficients de chaque composante sont sommés deux à deux, donc le nombre de bits nécessaires à l'expression de l'histogramme sera divisé par deux.

L'unité de base de la transformation de Haar se compose d'une opération de somme et d'une opération de différence. La sommation des coefficients revient à effectuer une moyenne des coefficients (filtres passe-bas), alors que leur différence correspond à un filtre passe-haut (le nombre de coefficients passe-bas est 16, les 240 restants sont les coefficients passe-haut). La différenciation est utile pour exprimer l'information contenue dans les résolutions fréquentielles plus fines.

A chaque exécution de la transformation, le nombre de niveaux d'expression de l'une des composantes H, S ou V est également divisé par deux. Il n'est pas possible de retrouver les 256 coefficients de base de l'histogramme à partir de 32 valeurs.

En revanche, deux histogrammes qui ne contiendraient pas le même nombre de coefficients peuvent toujours être comparés : il suffit pour cela d'effectuer la transformation de Haar sur l'histogramme qui contient le plus de coefficients jusqu'à ce qu'il en ait autant que le deuxième.

De même, lorsqu'une application ne nécessite que peu de coefficients, on peut transformer l'histogramme en notre possession pour ne lui envoyer que le nombre de coefficients adéquat.

Nombre de coefficients	Nombre de niveaux H	Nombre de niveaux S	Nombre de niveaux V
16	4	2	2
32	8	2	2
64	8	2	4
128	8	4	4
256	16	4	4

Tableau 3.4 : nombre de niveaux H, S et V en fonction du nombre de coefficients.

- **Représentation**

Le nombre de bins de l'histogramme et le nombre de bits de poids faibles ignorés sont représentés respectivement par les champs *NumberOfCoefficients* et *NumberOfBitplanesDiscarded*.

Le premier champ indique si l'une des valeurs (coefficients) suivantes est utilisée 16, 32, 64, 128 ou 256, alors que le deuxième champ spécifie le nombre de bits (vari entre 0 et 8) de poids faibles ignorés pour les coefficients.

Pour chaque coefficient, le signe doit être gardé et stocké, et représenté dans un tableau *CoefficientSign*.

Les valeurs des coefficients sont stockées dans un tableau de deux dimensions, *BitPlane*, ce qui permet de facilement ignorer les bits de poids faible afin de ne transmettre que les bits de poids forts lorsque nécessaire.

La représentation de descripteur *Scalable Color* est montrée dans le tableau suivant :

Champ	Nombre de bits	Description
NumberOfCoefficients	3	Spécifie le nombre valeurs (bins) de l'histogramme (16, 32, 64, 128 ou 256).
NumberOfBitplanesDiscarded	3	Spécifie le nombre de bits ignorés
CoefficientSign	1	Signe de chaque coefficient
BitPlane	variable	Valeurs des coefficients, sous forme de tableau.

Tableau 3.5 : représentation du descripteur Scalable Color.

Le descripteur *Scalable Color* est aussi représenté sous format XML comme suit :

```
<VisualDescriptor xsi:type="ScalableColorType"
numOfBitplanesDiscarded="0"
numOfCoeff="64">
<Coeff>-40 1 -12 63 -14 11 3 5 31 26 -5 9 -54 -2 12 9 -7 2 3 1 -3 5 2 -1 9
-2 1 1 -3 5 1 -4 3 3 1 3 3 2 -2 2 1 1 1 3 1 0 3 5 -9 3 2 0 -2 0 1 -3 0 0 0
-2 1 0 -3 -3</Coeff>
</VisualDescriptor>
```

Figure 3.13 : descripteur Scalable Color représenté sous format XML.

- **Correspondance**

Pour calculer la distance entre deux descripteurs de type *Scalable Color*, on utilise la norme L1, c'est-à-dire la somme des différences absolues (distance de Hamming), appliquée sur les coefficients  $C_i$  issus de la transformée de Haar, comme montré dans l'équation suivante :

$$D = \sum_i |C1i - C2i| \quad (6)$$

### III.7.5 Le descripteur Group Of Pictures or Group Of Frames (GOP/GOF) [17, 19]

- **Définition**

C'est une extension de descripteur *Scalable Color*, utilisé pour la représentation des caractéristiques couleurs d'un groupe d'images ou d'un groupe de trames dans un segment vidéo.

Le descripteur *GOP/GOF* permet de générer un histogramme couleur pour un segment vidéo ou un groupe d'images.

- **Extraction**

Les descripteurs *GOP/GOF* sont obtenus en agrégeant les histogrammes d'un groupe d'images ou de trames de vidéo, et en représentant les histogrammes agrégés en utilisant le descripteur SCD.

Trois valeurs sont définies pour calculer les bins de l'histogramme couleur agrégé pour toutes les séries d'images ou trames vidéo : la moyenne, la médiane et l'intersection. L'histogramme agrégé est ainsi entré en transformation de Haar pour construire la représentation SCD.

L'histogramme de moyenne est calculé en accumulant les histogrammes dans un groupe de trames ou d'images, ensuite en normalisant chaque bin accumulée par N, tel que N est le nombre de trames dans GOF ou bien le nombre d'images dans GOP. Un problème potentiel avec l'utilisation des moyennes simples pour calculer l'histogramme GOP/GOF est la sensibilité de l'opérateur moyen aux annexes.

L'histogramme de la médiane est obtenu en construisant pour chaque bin la liste croissante de N valeurs d'histogrammes de trames ou d'images au dessus de la longueur du GOF ou GOP. Et en assignant la médiane de cette liste au bin correspondant dans l'histogramme GOP/GOF.

L'histogramme de la médiane élimine les effets anormaux comme les changements de la lumière, l'occlusion, les recouvrements des textes, ainsi de suite, auxquels l'histogramme de moyenne est vulnérable.

L'histogramme d'intersection (Int\_Histogram) est obtenu en calculant la valeur minimale de chaque bin de trames ou d'images. Ensuite, en assignant cette valeur au bin d'histogramme résultant.

L'intersection trouve les couleurs communes minimales dans les trames ou les images.

- **Représentation**

La représentation du descripteur GOP ou GOF est identique à celle du SCD, à la seule différence de rajouter l'agrégation (*Aggregation*) comme un attribut supplémentaire.

Les trois différentes méthodes de l'agrégation sont représentées par cet attribut en utilisant deux bits. Ceci est suivi par le descripteur *Scalable Color* associé.

Champ	Nombre de bits	Description
NumberOfCoefficients	3	Spécifie le nombre valeurs (bins) de l'histogramme (16, 32, 64, 128 ou 256).
NumberOfBitplanesDiscarded	3	Spécifie le nombre de bits ignorés
CoefficientSign	1	Signe de chaque coefficient
BitPlane	variable	Valeurs des coefficients, sous forme de tableau.
Aggregation	2	Spécifie les trois différents types de l'agrégation

Tableau 3.6 : représentation du descripteur GOP or GOF.

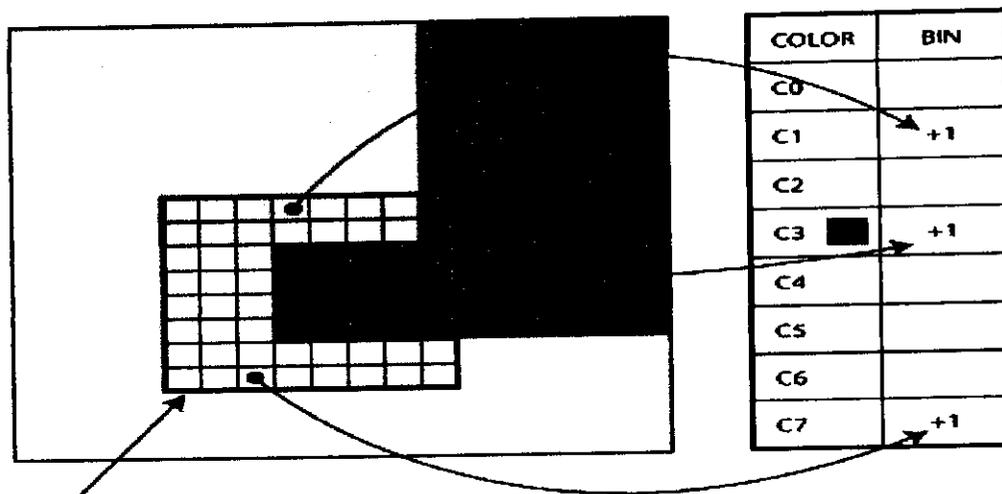
- **Correspondance**

Pour calculer la distance entre deux descripteurs de type *GOP/GOF*, on utilise le même principe que celui qu'on a utilisé pour le descripteur *Scalable Color*.

### III.7.6 Le descripteur Color Structure [17, 22]

Le descripteur *Color Structure* (CSD) est une généralisation de l'histogramme de couleur qui capture certaines caractéristiques de la distribution de la couleur dans une image. C'est un descripteur caractéristique couleur qui capture au même temps le contenu couleur (similaire à un histogramme couleur) et l'information de la structure de ce contenu.

Ce descripteur est utilisé généralement pour la recherche dans les images fixes, il exprime la structure couleur locale en utilisant un élément structurant (voir la **figure 3.14**).



Élément structurant

Figure 3.14 : élément structurant pour le descripteur Color Structure.

- Extraction

Le descripteur CSD compte combien de fois une couleur particulière est contenue dans l'élément structurant. Supposant les  $M$  couleurs quantifiées  $C_0, C_1, C_2, \dots, C_{M-1}$ , un histogramme structure couleur peut alors être défini par  $h(m)$  tel que  $m \in \{0, 1, \dots, M-1\}$  où la valeur dans chaque bin d'histogramme est égale au nombre d'éléments structurants dans l'image contenant un ou plusieurs pixels avec la couleur  $C_m$ .

Le nombre d'éléments structurants est toujours 64, la distance entre les points augmente avec la taille de l'image, comme montré dans la figure 3.15. Cette méthode est équivalente à un sous-échantillonnage de l'image par la puissance de 2 et alors l'utilisation d'un élément structurant de 8x8 pixels.

Le facteur de sous-échantillonnage est donné par  $K = 2^p$ , tel que:

$$P = \max \{0, \text{rond}(0.5 * \log_2(\text{largeur} * \text{hauteur}) - 8)\}. \quad (7)$$

$$E = 8 * K.$$

$E \times E$  est l'affluence spatiale de l'élément structurant.

Comme montré dans la figure 3.15, pour les images plus petites que 320 x 240 pixels, un élément structure 8x8 sans sous-échantillonnage est utilisé, et pour les images de taille 640 x 480 ( $p = 1, K = 2, E = 16$ ), l'élément structurant est 16x16 et le sous-échantillonnage est 2x2.

L'élément structurant de taille 8x8 est appliqué pour une image sous-échantillonnée.

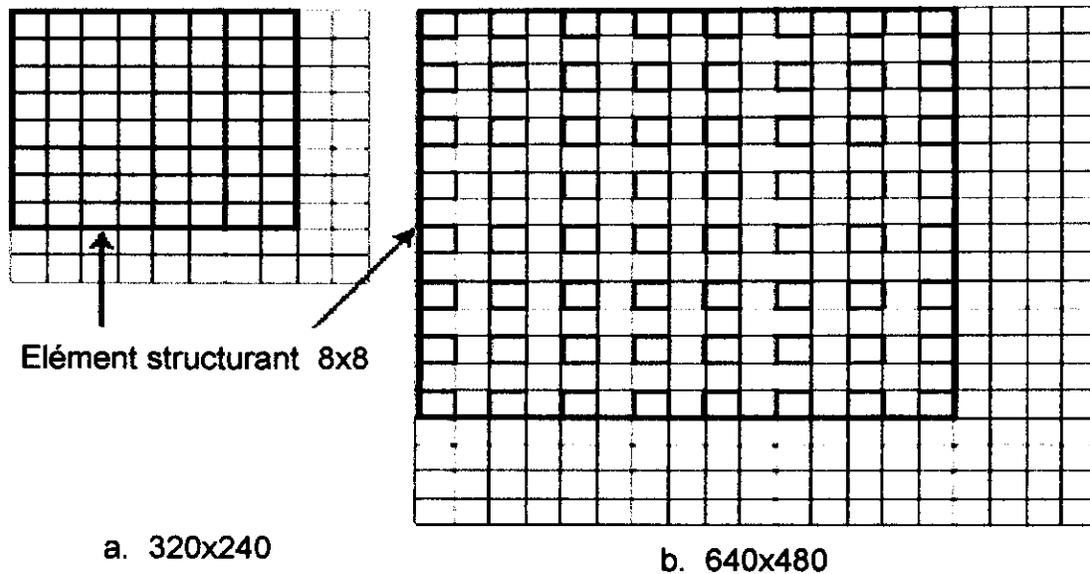


Figure 3.15 : éléments structurants avec des images de différentes tailles.

Le descripteur CSD utilise l'espace couleur HMMD, et définit en utilisant 4 points de quantification de l'espace couleur (4 Color Space Quantization Operating Points). La quantification, montrée dans le **tableau 3.7** est faite suivant les étapes suivantes :

- ✓ Premièrement, l'espace HMMD est divisé en 5 sous-espaces (0, 1, 2, 3 et 4). La division du sous-espace est exécutée sur la valeur Diff de l'espace HMMD par les intervalles suivants : [0, 6], [6, 20], [20, 60], [60, 110], et [110, 255].
- ✓ Deuxièmement, chaque sous-espace couleur est uniformément quantifié sur les axes de la teinte (Hue) et la somme (Sum) où le nombre de niveaux de quantification pour les paramètres Hue et Sum est montré dans le tableau.
- ✓ Dans la dernière étape, les valeurs qu'il contient l'histogramme de structure couleur sont normalisées en 8 bits, alors l'intervalle de l'histogramme est convertit de [0, 1] à 0,255.

Nombre de cellules	256		128		64		32	
	Hue	Sum	Hue	Sum	Hue	Sum	Hue	Sum
0	1	32	1	16	1	8	1	8
1	4	8	4	4	4	4	4	4
2	16	4	8	4	4	4		
3	16	4	8	4	8	2	4	1
4	16	4	8	4	8	1	4	1

Tableau 3.7 : quantification de l'espace HMMD pour le CSD.

- Représentation

Le descripteur Structure Color est présenté sous format XML comme suit :

```
<VisualDescriptor xsi:type="ColorStructureType" colorQuant="2">
<Values> 0 0 0 0 0 0 0 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 34 85 170 103 0 0 0 0 7 5
0 0 0 3 2 0 43 110 146 150 0 7 29 28 32 52 32 4 54 64 45 18 8 28 53 139 137 93
44 11</Values>
</VisualDescriptor>
```

Figure 3.16 : représentation de descripteur Color Structure sous format XML.

## CONCLUSION

Dans ce chapitre, nous avons présenté le standard MPEG-7, les différents descripteurs visuels définis par ce standard. L'attribut visuel qui nous intéresse est la couleur donc, nous avons détaillé surtout les descripteurs couleurs.

Dans la suite de notre travail, nous considérons les différents descripteurs couleurs de MPEG-7 pour l'indexation et la recherche par similarités de contenus d'images.

# Chapitre4

## *Conception et mise en œuvre*

## INTRODUCTION

Dans notre travail, nous nous sommes intéressés particulièrement à l'attribut couleur pour résoudre la problématique de recherche d'image par le contenu. Nous présentons ici notre conception d'un système d'indexation et de recherche d'images par similarités, basées sur les descripteurs de contenu MPEG-7, notamment les descripteurs couleurs. L'architecture générale d'un système d'indexation et de recherche d'image par le contenu est illustrée dans la **figure 4.1**.

Notre travail est composé de deux phases différentes : la première est la phase d'indexation, la deuxième phase est la recherche des images similaires.

### ✓ Phase d'indexation

On a mentionné dans le chapitre précédent que le standard MPEG-7 définit des descripteurs visuels permettant de décrire les différentes caractéristiques des contenus d'images. L'objet de notre travail est l'attribut couleur.

Les descripteurs choisis dans notre projet pour la recherche d'images similaires sont : le descripteur Color Layout (CLD), le descripteur Scalable Color (SCD) et le descripteur Dominant Color (DCD). L'extraction de ces trois descripteurs est faite en utilisant l'outil MPEG-7[24]. Cet outil définit des algorithmes décrivant les différentes étapes d'extraction de chaque descripteur.

### ✓ Phase de recherche

Notre système de recherche s'appuie sur la sélection d'une image quelconque, puis la recherche des images qui se rapprochent de cette *image requête*. Le système doit donc chercher les images qui ont des index similaires à l'index de l'image requête. Les images résultantes sont alors affichées par un ordre d'affichage en utilisant les distances de similarités.

### IV.1 ARCHITECTURE GÉNÉRALE DU SYSTÈME

La figure suivante représente un modèle de système pour l'indexation et la recherche d'images par le contenu.

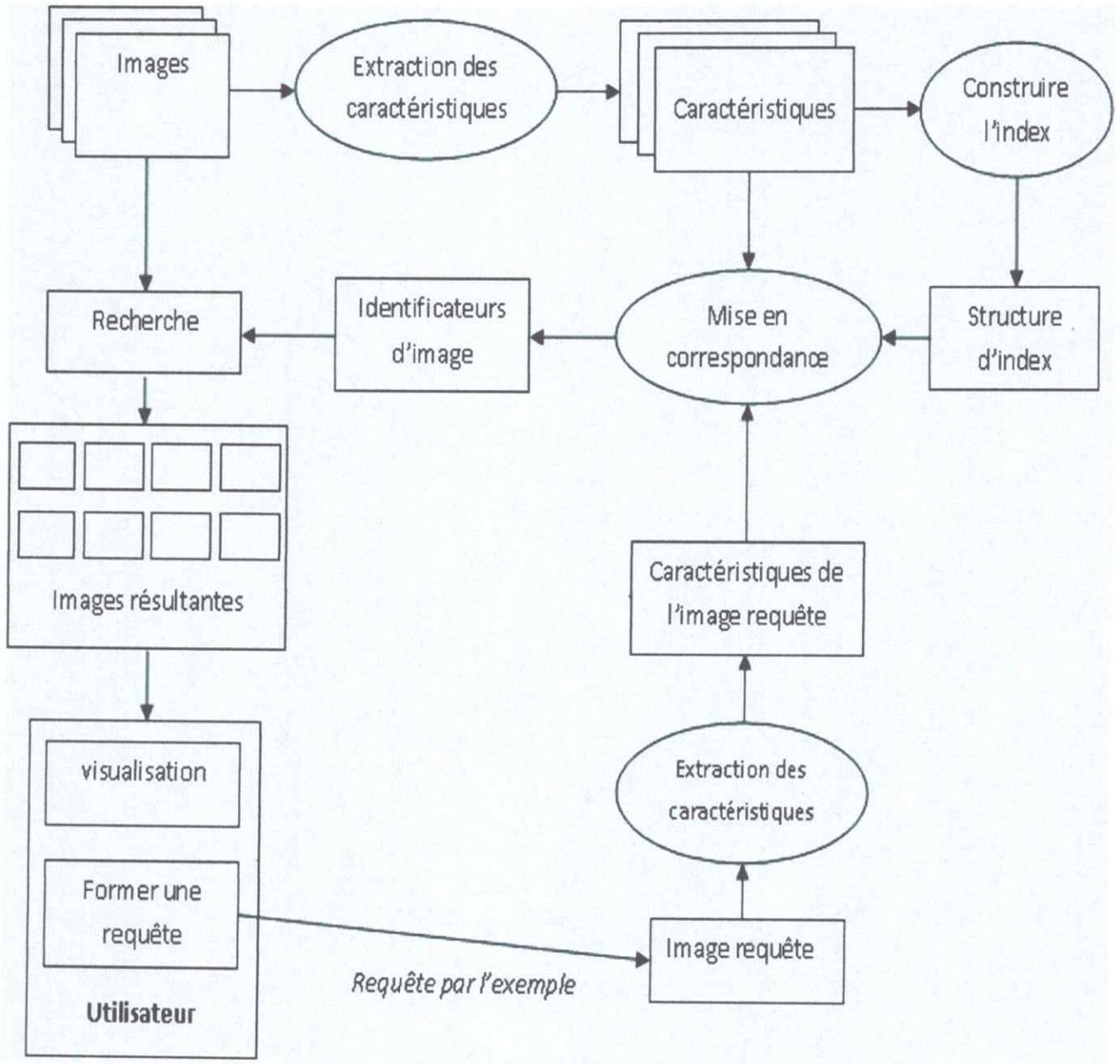


Figure 4.1 : Architecture générale d'un système de recherche d'image par le contenu couleur

## IV.2 DESCRIPTION DE L'ARCHITECTURE

Nous proposons pour l'extraction des caractéristiques, ainsi que la recherche d'images par similarité, de considérer trois descripteurs couleurs.

### IV.2.1 Module d'indexation

Pour l'indexation, on va décrire les différentes procédures utilisées par notre système permettant d'extraire les trois descripteurs couleurs CLD, SCD ainsi que le descripteur DCD.

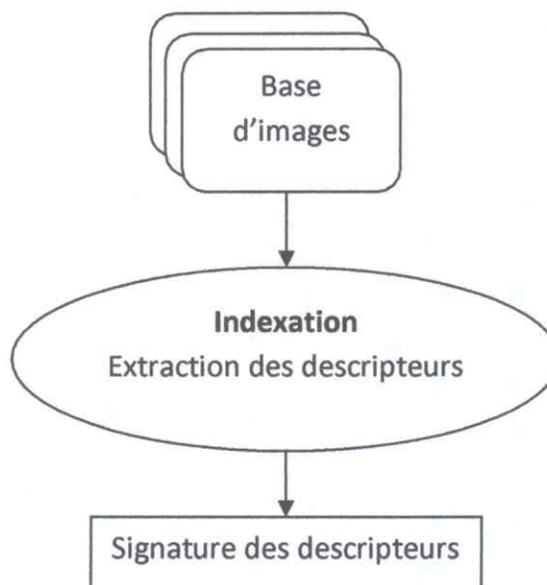


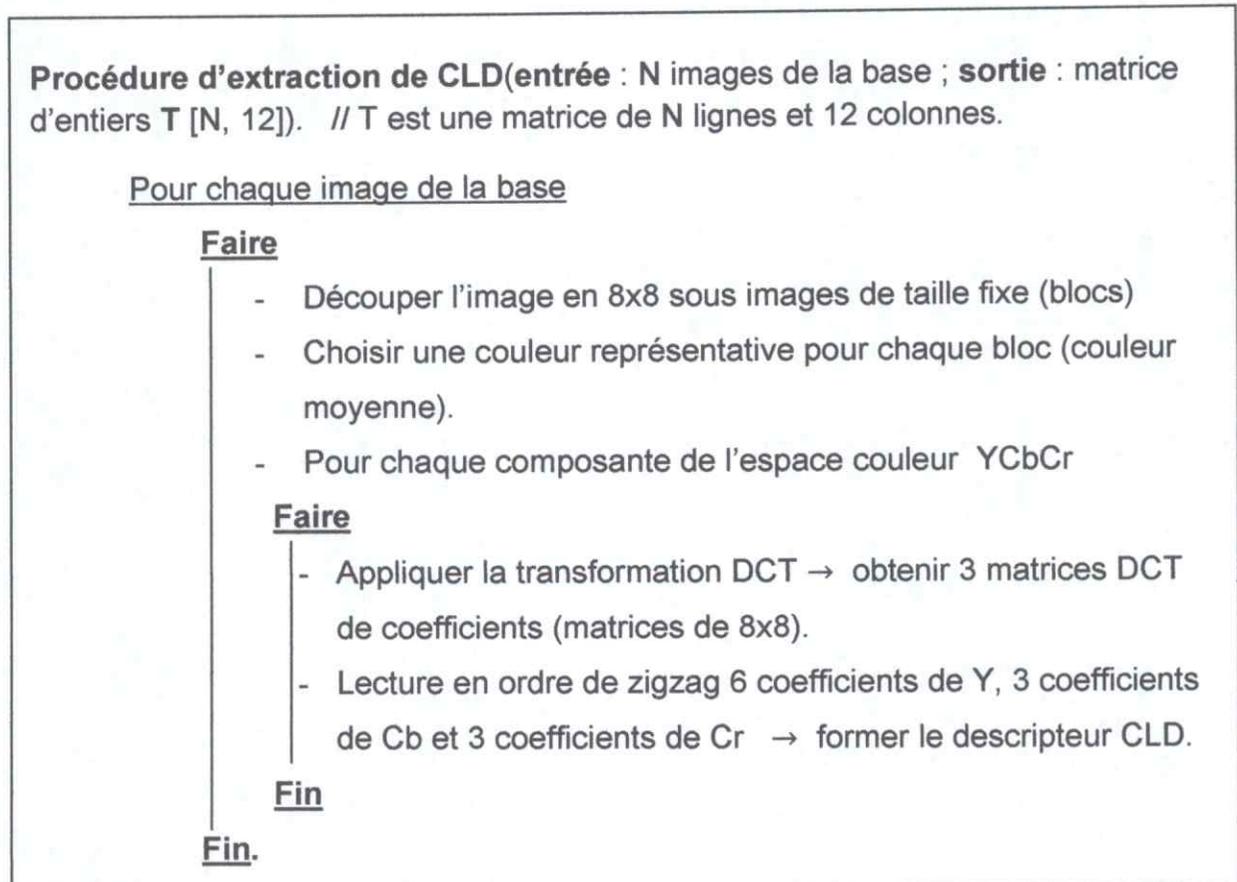
Figure 4.2 : Module d'indexation

#### IV.2.1.1 Extraction de descripteur Color Layout

L'espace couleur utilisé pour l'extraction de CLD est YCbCr, et le nombre de coefficients extraits pour chaque composante est variable. Ce nombre peut prendre une des valeurs suivantes : 1, 3, 6, 10, 15, 21, 28, et 64.

Dans notre cas, ce descripteur utilise 12 coefficients, 6 pour la luminance (Y), 3 pour la chrominance bleue et la chrominance rouge (Cb et Cr).

Ces coefficients sont extraits de la manière suivante :



Le résultat de cette procédure est une matrice T de N lignes et 12 colonnes tel que :

- $T [i, j] = k$  avec :  $i = 1..N$  et  $j = 1..12$  ; où  $i =$  l'image « i » de la base,  $j =$  le  $j^{\text{ème}}$  coefficient Color Layout et  $k =$  instance du coefficient « j » pour l'image « i ».

Les différentes étapes de l'extraction de CLD sont illustrées dans la figure suivante :



Figure 4.3 : processus d'extraction de CLD (étapes).

#### IV.2.1.2 Extraction de descripteur Scalable Color

Scalable Color utilise 256 paramètres extraits de la manière suivante :

**Procédure d'extraction de SCD (entrée : N images de la base ; sortie : matrice d'entiers T1 [N, 256]). // T1 est une matrice de N lignes et 256 colonnes.**

Pour chaque image de la base

Faire

- Extraire dans l'espace couleur HSV un histogramme de couleurs (16 valeurs pour H, chacune avec 4 valeurs pour S et 4 pour V) → 256 valeurs.
- Quantifier non linéairement les 256 valeurs de l'histogramme en 4 bits.
- Appliquer la transformation de Haar sur les valeurs quantifiées en 4 bits.
- Quantifier linéairement les valeurs résultantes de la transformation de Haar.
- Extraire les 256 coefficients (16 coefficients passe-bas, 240 coefficients passe-haut).

Fin.

Le résultat de cette procédure est une matrice T1 de N lignes et 256 colonnes tel que :

- $T[i, j] = k$  avec :  $i = 1..N$  et  $j = 1..256$  ; où  $i =$  l'image «  $i$  » de la base,  $j =$  le  $j^{\text{ème}}$  coefficient Scalable Color et  $k =$  valeur du coefficient «  $j$  » pour l'image «  $i$  ».

#### IV.2.1.3 Extraction de descripteur Dominant Color

Contrairement aux descripteurs CLD et SCD, le nombre de coefficients utilisés par le descripteur DCD est variable d'une image à l'autre selon le nombre de couleurs dominantes qu'elle contient.



Dans notre travail, on ne prend pas en compte les paramètres de la variance et la cohérence spatiale. La procédure d'extraction est la suivante :

- Au début, la procédure est initialisée avec un seul cluster comprenant l'ensemble des pixels de l'image et une couleur représentative de ce cluster est calculée comme un centre de masse. A chaque itération, l'algorithme recalcule les clusters de couleurs et calcule une fonction de distorsion.
- Ceci se fait jusqu'à ce qu'un critère d'arrêt est atteint (distorsion minimale ou nombre maximal d'itérations).
- Des perturbations sont ensuite ajoutées aux centres de masse des clusters ayant les distorsions les plus grandes afin de les faire diminuer. L'algorithme s'arrête alors lorsque la distorsion maximale est inférieure à un certain seuil, ou que le nombre maximal de clusters est atteint.
- Le pourcentage  $p_i$  de pixels de l'image appartenant à chacune des couleurs dominantes est ensuite calculé, le résultat est uniformément quantifié sur 5 bits. Les valeurs couleurs sont ensuite quantifiées sur 1 à 12 bits en fonction de l'espace couleur et les descripteurs de quantification des couleurs associés.

Le nombre de paramètres extraits est variable d'une image à l'autre, par exemple pour une image contenant 7 couleurs dominantes, le nombre de paramètres extraits est 29.

Le premier paramètre extrait représente le nombre de couleurs dominantes, et pour chaque couleur dominante, un paramètre représente le pourcentage et trois paramètres représentent les valeurs de cette couleur.

#### **IV.2.2 Modèle de représentation des données**

Un modèle de représentation des données d'images est nécessaire pour stocker, interroger des images et les intégrer dans les SGBD classiques. Nous proposons un modèle qui permet de stocker l'ensemble d'images ainsi que les différentes caractéristiques extraites à partir de chaque image de la base.

1. **Base d'images** : contient toutes les images.
2. **Base d'index** : le résultat d'indexation pour chaque descripteur est mis dans un document texte, où chaque ligne correspond aux paramètres extraits de chaque image de la base.
3. **Base de données** : nous avons construit une base de données pour stocker les données (paramètres) de la base d'index.

Notre base de données contient les trois tables suivantes :

- Une **table layout** est définie pour représenter le descripteur Color Layout, elle présente la structure suivante : layout (num\_image, nom\_image, chem\_image, P).

*num\_image* pour identifier l'image, il représente le numéro séquentiel de l'image dans la base.

*nom\_image* représente le nom de l'image.

*chem\_image* représente l'emplacement de l'image.

*P* vecteur de caractéristiques de l'image, il contient 12 champs représentant les coefficients de Color Layout extraits de chaque image.

- Une **table scalable** est définie pour représenter le descripteur Scalable Color, elle présente la structure suivante : scalable (num\_image, nom\_image, chem\_image, F) tel que :

Les trois champs *num\_image*, *nom\_image* et *chem\_image* ont le même rôle que dans la table layout.

*F* vecteur de caractéristiques de l'image, il contient 256 champs représentant les coefficients Scalable Color extraits de chaque image de la base.

- Une **table dominant** est définie aussi pour représenter le descripteur Dominant Color, elle présente la structure suivante : dominant (num\_image, nom\_image, chem\_image, nbre\_domin, T) tel que :

Les trois champs *num\_image*, *nom\_image* et *chem\_image* ont le même rôle que dans les tables layout et scalable.

*nbre\_domin* permet de représenter le nombre de couleurs dominantes dans chaque image de la base.

*T* un vecteur de caractéristiques de l'image, il contient les champs représentant les pourcentages de chaque couleur dominante d'une part, et les champs représentant les variables couleurs de chaque couleur dominante d'autre part.

### IV.3 Module de recherche

Les résultats obtenus dans l'indexation nous permettent de passer l'étape suivante, la recherche des images similaires. Le module de recherche est schématisé dans la figure suivante :

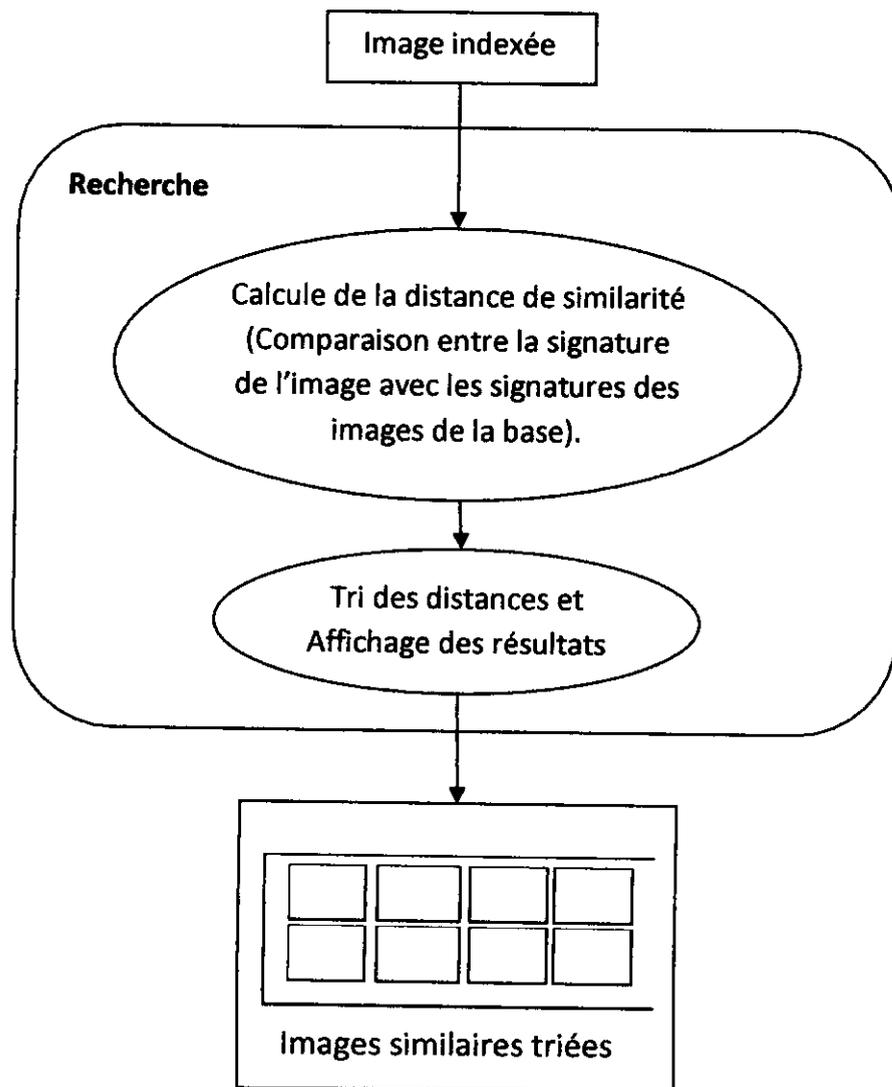


Figure 4.4 : Module de Recherche.

Notre système de recherche opère avec deux méthodes différentes :

- Une recherche unique qui permet la recherche des images similaires en utilisant un seul descripteur.
- Une recherche multiple qui permet de combiner entre les descripteurs.

### IV.3.1 Recherche unique

Une recherche par similarité basée sur un descripteur s'opère par comparaison. On calcule la distance de similarité entre la signature de l'image sélectionnée (index de l'image requête) et les signatures des autres images de la base.

Étant donnée une image requête  $I_q$ , afin de calculer la distance de similarité entre la signature de  $I_q$  et la signature d'une image quelconque de la base ( $I$ ), nous utilisons les distances décrites dans le chapitre précédent.

- Pour le descripteur Scalable Color, la distance utilisée est celle de *Hamming*, elle est calculée par la formule suivante :

$$d(S_1, S_2) = \sum_{i=1}^N |P1i - P2i|.$$

- Pour Color Layout, on a utilisé la distance euclidienne calculée par la formule suivante :

$$d(S_1, S_2) = \sum_{i=1}^N \sqrt{(P1i - P2i)^2}.$$

Avec :  $S_1$  : signature (index) de l'image requête.

$S_2$  : signature de l'image de la base.

$P1i$  : le  $i^{\text{ème}}$  élément (paramètre) extrait de l'image requête.

$P2i$  : le  $i^{\text{ème}}$  élément (paramètre) extrait de l'image de la base.

- Pour le descripteur Dominant Color, nous avons calculé la distance de similarité en utilisant la formule proposée par MPEG-7 (équation 3). Dans notre cas, nous avons choisis  $T_d = 15$  et  $\alpha = 1,2$ .

Les différentes procédures de recherche se basant sur les descripteurs CLD, SCD et DCD sont décrites comme suit :

#### • Procédure de recherche pour Scalable Color

**Procédure de la recherche de SCD (entrée : matrice d'entiers T1 [N, 256] ; Sortie : tableau de valeurs réels Dist [N])**

Sélectionner une image de la base (image requête q)

Pour chaque image de la base (pour l = 1 à N)

**Faire**

- Calculer  $Dist [l] = \sum_{i=1}^{256} | Coef1[i] - Coef2[i] |$  ;  
// la distance de similarité entre l'image requête et l'image « i » ;
- Mettre les distances dans un ordre croissant
- Afficher les images selon l'ordre de distances ;

**Fin**

- Coef1[i] et Coef2[i] sont respectivement les valeurs des coefficients d'ordre « i » de l'image requête et l'image de la base.
- Dist[l] est la distance de similarité entre l'image de la base « l » et notre image requête.

• Procédure de recherche Color Layout

**Procédure de la recherche de CLD (entrée : matrice d'entiers T [N, 12] ; Sortie : tableau de valeurs réels Dist [N])**

Sélectionner une image de la base (image requête q)

Pour chaque image de la base (pour l = 1 à N)

**Faire**

$$Dist[I] = \sqrt{\sum_{i=1}^6 W_{yi} \times (Coef_{Y1[i]} - Coef_{Y2[i]})^2} + \sqrt{\sum_{i=1}^6 W_{Cb_i} \times (Coef_{Cb1[i]} - Coef_{Cb2[i]})^2} + \sqrt{\sum_{i=1}^3 W_{Cr_i} \times (Coef_{Cr1[i]} - Coef_{Cr2[i]})^2}$$

// la distance de similarité entre l'image requête et l'image « i » ;

Mettre les distances dans un ordre croissant

Afficher les images selon l'ordre de distances ;

**Fin**

-  $w_i$  dénote les valeurs poids associés à chaque coefficient, les plus grands poids sont associés aux coefficients de plus grande fréquence selon l'ordre en zigzag. Un exemple de valeurs poids est affiché dans le tableau suivant :

Coefficient	Y1	Y2	Y3	Y4	Y5	Y6	Cb1	Cb2	Cb3	Cr1	Cr2	Cr3
poids	3	3	3	1	1	1	2	1	1	4	2	2

Tableau 4.1 : exemple de valeurs poids des coefficients CLD.

- Coef\_Y1 [i] et Coef\_Y2[i] représentent respectivement la valeur du coefficient d'ordre « i » de la composante Y pour l'image requête et l'image de la base.

- Coef\_Cb1[i] et Coef\_Cb2[i] sont respectivement les valeurs du coefficient d'ordre « i » de la composante Cb pour l'image requête et l'image de la base.

- Coef\_Cr1[i] et Coef\_Cr2[i] sont respectivement les valeurs du coefficient d'ordre « i » de la composante Cr pour l'image requête et l'image de la base.

• Procédure de recherche pour Dominant Color

**Procédure de la recherche de DCD** (entrée : matrice d'entiers T2[N, P] ; Sortie : tableau de valeurs réels Dist [N])

Sélectionner une image de la base (image requête q)

Pour chaque image de la base (pour l = 1 à N)

**Faire**

Pour i = 1 à N1 faire

Pour j = 1 à N2 faire

Calculer  $d(1i, 2j) = \sqrt{(C1i - C2j)^2}$

Si  $d(1i, 2j) < Td$  alors

$a(1i, 2j) = (1 - d(1i, 2j)) / (\alpha * Td)$  ;

sinon  $a(1i, 2j) = 0$  ;

//calcul de distance de similarité

$Dist^2 [l] = \sum_{i=1}^{N1} (P1i)^2 + \sum_{j=1}^{N2} (P2j)^2 - \sum_{i=1}^{N1} \sum_{j=1}^{N2} 2 * a(1i, 2j) * P1i * P2j$  ;

Mettre les distances dans un ordre croissant

Afficher les images selon l'ordre de distances ;

**Fin**

- C1i : la i<sup>ème</sup> couleur dominante dans l'image requête.
- C2j : la j<sup>ème</sup> couleur dominante dans l'image de la base.
- N1 et N2 représentent respectivement le nombre de couleurs dominantes pour l'image requête (q) et l'image de la base (l).
- d(1i, 2j) est la distance euclidienne entre une couleur de l'image requête et l'autre de l'image de la base.
- a(1i, 2j) est le coefficient de similarité entre deux couleurs.
- Dist<sup>2</sup>[l] est la distance de similarité entre l'image de la base « l » et notre image requête.

Pour trier la liste d'images similaires, il est utile d'utiliser un algorithme de tri qui permet d'ordonner les images résultantes selon leur distance avec l'image requête. Cet algorithme classe les distances de similarité dans un ordre croissant, où la plus petite distance correspond à l'image la plus proche de notre image requête.

### IV.3.2 Recherche multiple (Combinaison des descripteurs)

L'idée de la combinaison entre les descripteurs est de trouver un processus de recherche qui utilise les caractéristiques complémentaires de chaque descripteur d'une manière appropriée, ce qui permet d'avoir les meilleurs résultats pour la recherche.

Pour faire la combinaison entre les descripteurs couleurs, on peut par exemple utiliser les opérateurs suivants : *addition*, *multiplication*, *minimum* et *maximum*.

L'opérateur *maximum* donne le nombre maximum d'images corrects obtenues entre les descripteurs, *minimum* donne le nombre d'image minimal.

Dans notre travail, nous avons conçu la combinaison en utilisant deux méthodes différentes :

#### Première méthode (calcul de la distance moyenne)

Dans ce cas, le principe est de calculer la distance moyenne entre les descripteurs à combiner de la manière suivante :

- 1- Calculer les distances de similarité pour chaque descripteur ( $d_i$ ).
- 2- Calculer la distance moyenne entre les descripteurs à combiner ( $D$ ).

La distance moyenne est calculée de la manière suivante :  $D = \sum w_i * d_i$  ; tel que :  $w_i$  sont des poids associés à la distance de similarité pour chaque descripteur.

#### Deuxième méthode (filtrage)

Cette méthode de combinaison est complètement différente par rapport à la méthode précédente, elle a pour principe :

Appliquer l'algorithme de recherche d'un des descripteurs, puis appliquer l'algorithme de recherche du deuxième descripteur sur les  $n$  images les plus proches obtenu par le premier descripteur (dans notre travail, nous avons choisi  $n = 40$ ).

## CONCLUSION

Dans ce chapitre, nous avons présenté la conception et la mise en œuvre de notre système d'indexation et de la recherche d'image en utilisant les trois descripteurs couleurs : Color Layout, Scalable Color et Dominant Color.

La recherche des images similaires à une image requête se résume alors à calculer la distance entre l'index de l'image requête et ceux des images de la base. Les plus petites distances se correspondent aux images les plus similaires.

# Chapitre 5

## *Réalisation et Résultats Expérimentaux*

## INTRODUCTION

Dans ce chapitre, nous allons décrire l'implémentation de notre logiciel d'indexation et de recherche d'images basées sur les descripteurs couleurs. Nous proposons par la suite une étude des divers tests, ainsi que les résultats obtenus.

Pour évaluer et valider notre système de recherche d'images, nous avons utilisé une base d'images disponible sur Internet. La base de J.Z. Wang contient 1000 images divisées en 10 classes contenant chacune 100 images (Wang et al., 2001) [24]. Certaines d'entre elles sont reprises ci-dessous.



Figure 5.1 : exemples d'images.

Le logiciel que nous avons conçu est un logiciel de recherche d'images dont la fenêtre principale (fenêtre de recherche) est représentée ci-dessous.

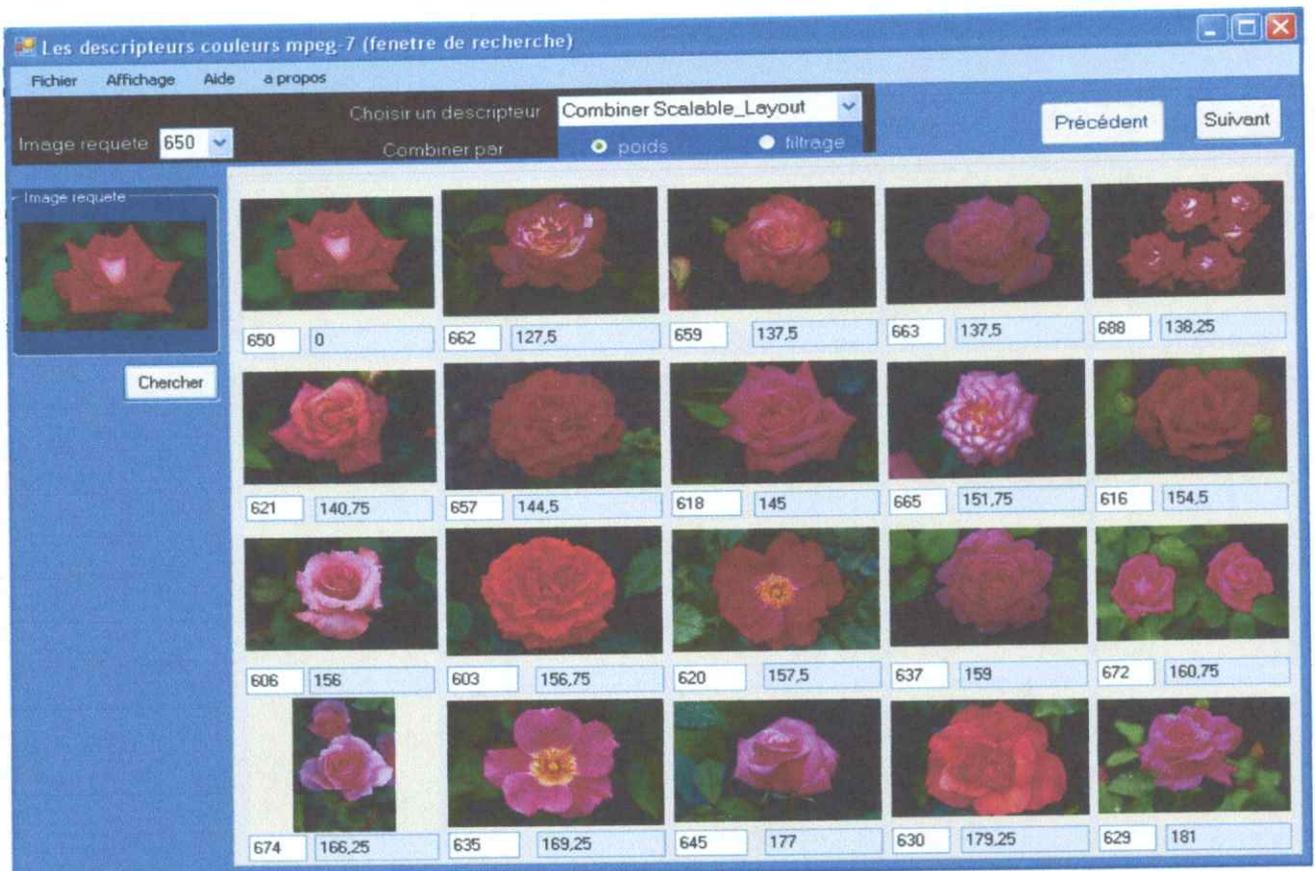


Figure 5.2 : fenêtre de recherche

## V.1 RÉSULTATS D'INDEXATION

Pour comparer les résultats d'indexation pour les trois descripteurs CLD, SCD et DCD, nous avons considéré trois images de notre base (image 4, image 7 et image 9). Chaque descripteur, pour les trois images considérées, sont présentés ci après.

### V.1.1 Descripteur Color Layout (CLD)

Les différents paramètres obtenus pour les trois images lors de l'indexation du descripteur CLD sont :

**Image4** : 15, 27, 51, 17, 24, 15, 19, 20, 16, 19, 16, 10.

**Image7** : 13, 28, 52, 16, 27, 23, 15, 14, 15, 21, 18, 8.

**Image9** : 14, 37, 27, 19, 26, 16, 19, 13, 16, 24, 16, 12.

Les 6 premiers paramètres de chaque image représentent les coefficients Y (entre Y1 et Y6), les 6 qui restent représentent respectivement 3 coefficients de Cb et 3 coefficients de Cr.

### V.1.2 Descripteur Scalable Color (SCD)

Les 256 paramètres résultants lors de l'indexation de descripteur SCD, sont affichés pour chaque image ci après :

**Image4** : 125, 20, 111, 89, 0, 9, -1, 21, -16, -22, -22, 6, 9, 12, -9, 24, -6, -4, 3, 3, -11, 3, -3, 21, -9, 0, -10, 10, -10, 5, 1, 24, -1, -3, -3, -3, 1, 1, 0, 2, -2, -6, -8, -4, -4, -6, -1, 5, -1, 1, 0, 2, 15, 3, 3, 2, 0, 1, 0, -2, 1, 0, -3, -1, -1, 0, 0, 0, 0, -1, -2, 3, 1, 2, 3, 2, 2, 2, 1, 4, 1, -1, -3, -1, 1, -3, -4, 3, -1, 1, 2, 1, 0, 3, 2, 1, 1, 0, -2, 1, -1, -3, -2, 2, 0, 2, 2, 1, 3, 3, 1, 3, -2, -1, 0, 1, 3, -4, -3, 0, -1, 0, 1, 0, 1, 0, -1, 6, -1, 0, 0, -1, 1, -1, 0, 0, 0, 0, -1, 1, -1, 0, 0, -1, -1, 0, -2, 1, 0, 0, 0, -1, 0, 1, 0, 0, 0, 1, 4, 0, 0, 0, 0, 1, 0, 1, 0, -3, -1, -1, -1, 1, -1, 0, 1, -3, 0, 0, -1, 3, 0, 1, 0, -2, 0, 0, 0, 0, 0, 1, 4, -1, 1, 0, 0, 0, 0, 0, 0, -2, -1, 0, -1, 1, 0, 1, 1, -3, -1, 1, -1, 0, 1, 1, 3, -2, -1, 0, 0, 0, 1, 0, 7, -2, 0, 1, 0, 3, 0, 0, 0, 0, -1, 1, 0, 3, 0, 0, 0, 0, -1, 0, 0, 0, 1, 0, 4, 0, -1, 0, 0, 0, 1, 0, 7.

**Image7** : 136, 20, 127, 52, -2, 10, 7, 17, 6, -10, -8, 23, 15, 9, -1, 17, -7, 7, -3, 7, -9, 7, -13, 11, -15, 7, -7, 12, -11, 5, -14, 23, 2, -3, -3, -4, -1, -1, 0, 4, -1, -3, -4, -4, -3, -5, -2, 2, -5, -2, 2, 1, 6, -2, -3, -4, 0, 4, 0, -2, 4, 0, -3, 0, 0, 0, 0, 1, 0, 0, 2, 2, 1, 3, 2, 0, 3, 4, 3, 2, -1, -1, -2, -1, -3, -4, 0, 3, 1, 2, 1, 3, 3, 2, 1, 3, -2, 0, -2, -1, -3, -4, 1, 4, 4, 2, 1, 1, 3, 1, 2, -2, 0, 0, -1, -1, -4, -1, -1, 1, 3, 1, 0, -2, 0, -1, 8, -1, -1, 0, 0, 0, 0, 0, 0, 0, -1, 0, 0, -1, 0, 0, 1, -1, 0, 0, 0, -1, -1, 1, -1, 0, 0, 1, 0, 1, 0, 1, 2, -2, 0, -1, 1, 1, 0, 0, 0, -3, 0, 0, -1, 0, -1, 1, 0, -2, 0, 0, -1, 0, 0, 1, 1, -1, 0, 0, 0, 0, 1, 3, -2, 0, 0, 1, 1, 0, 1, 1, -3, 0, 0, -1, 2, -1, 2, 1, -1, 0, 1, -1, 2, 1, 1, 3, -1, -2, 0, 0, -2, 1, 0, 6, -2, 0, 1, 2, -1, 0, 0, 0, 0, 1, 0, 3, 0, 0, 1, 2, -1, 0, 0, 3, 1, 0, 3, 0, -3, 0, 0, -1, 1, 0, 7.

**Image 9** : -15, 54, 11, -3, -43, -15, -12, 16, -31, 1, 18, 14, -3, -3, -16, 17, 7, 0, -3, 9, -1, 5, 0, 0, -2, 7, -15, 3, -3, 5, -2, -4, -3, -3, -3, 3, 1, -1, 0, -6, 1, 1, 1, 3, 1, 2, 4, 5, 3, 2, 2, 1, -15, 2, 4, 2, 0, 0, 0, -2, -2, 0, -3, -3, -1, -1, -2, -6, -1, -1, -2, 2, 1, 1, 2, 1, 2, 2, 2, 1, -3, -1, -3, -7, -1, -3, 0, 0, 2, 1, 2, 1, 3, 3, 2, 0, -2, -2, -3, -5, 3, -1, 2, 1, 2, 2, 2, 1, 0, 3, 1, 0, -5, -3, -3, -4, -3, -4, 0, -3, 1, 1, 1, 0, 1, 0, -1, 1, -1, 1, -2, -2, 1, -1, 0, 0, 1, 0, -2, 0, 0, 0, 0, 1, 0, 0, -1, 0, 0, 1, -1, 0, 0, 0, 0, 0, 0, 1, 0, -1, -1, -1, 0, 1, -1, -1, -1, 0, 0, 0, -1, 1, 0, 0, -1, 0, 0, 0, -1, 0, 0, 1, -2, 0, 0, 0, 0, 0, 0, 1, -1, 0, 0, -1, 2, 1, -1, 0, -1, 0, 0, 0, -1, 3, -2, 1, -1, 0, 0, 1, -1, 3, 1, 1, -2, 0, 0, 0, 0, 0, 1, 0, -1, -1, 0, 0, 1, 0, 0, -1, -2, 0, 0, 1, 0, 0, 0, 0, -1, 0, 0, 0, 0, 0, 1, 0, -1, 0, 0, 0, 0, 0, 1, 0, 1.

### V.1.3 Descripteur Dominant Color (DCD)

**Image 4** : 8, 12, 1, 1, 1, 20, 25, 1, 4, 8, 12, 18, 24, 11, 16, 10, 10, 3, 27, 22, 6, 15, 20, 20, 21, 9, 20, 5, 6, 17, 6, 7, 8.

**Image 7** : 8, 16, 1, 1, 1, 19, 25, 1, 5, 14, 19, 23, 28, 12, 13, 5, 5, 10, 21, 4, 6, 5, 22, 18, 19, 19, 7, 7, 8, 0, 27, 22, 11.

**Image9** : 4, 50, 5, 5, 5, 31, 13, 15, 20, 8, 22, 23, 23, 9, 11, 12, 12.

Le premier paramètre représente le nombre de couleurs dominantes extraits à partir de l'image, par exemple l'image4 contient 8 couleurs dominantes.

Les valeurs affichées en bleu représentent les pourcentages des couleurs dominantes, et les 3 valeurs après les pourcentages représentent l'index de la couleur.

### Remarque

En comparant les paramètres résultants de l'extraction pour chaque image, on remarque que pour les deux premières images, les paramètres extraits sont proches l'un de l'autre, ce qui signifie que la distance de similarité pour chaque descripteur sera petite.

Par contre, pour l'image9, les paramètres extraits sont complètement différents par rapport au deux premières images, donc la distance de similarité sera grande.

## V.2 CRITÈRES D'ÉVALUATION

Plusieurs critères sont pris en compte pour évaluer et comparer la performance les différents descripteurs. Il existe deux méthodes complémentaires pour évaluer les résultats de la recherche automatique :

- Une première méthode repose sur des critères numériques d'évaluation qui permettent de tracer des courbes de comparaison entre descripteurs et de réaliser des statistiques.
- La deuxième méthode (c'est la méthode utilisée dans notre test) repose sur un jugement visuel de la qualité des résultats, en utilisant une interface graphique.

## V.3 RÉSULTATS DE LA RECHERCHE

Pour afficher les images similaires par l'attribut couleur à une image requête (image de référence), on utilise une interface graphique (fenêtre de recherche).

Les images, ainsi que les distances entre les images sont affichées par ordre de similarité décroissant

On va spécifier ci-dessous les différents résultats obtenus pour chaque descripteur en comparant ainsi ces résultats, ce qui nous permet de savoir le degré de la performance pour chaque descripteur.

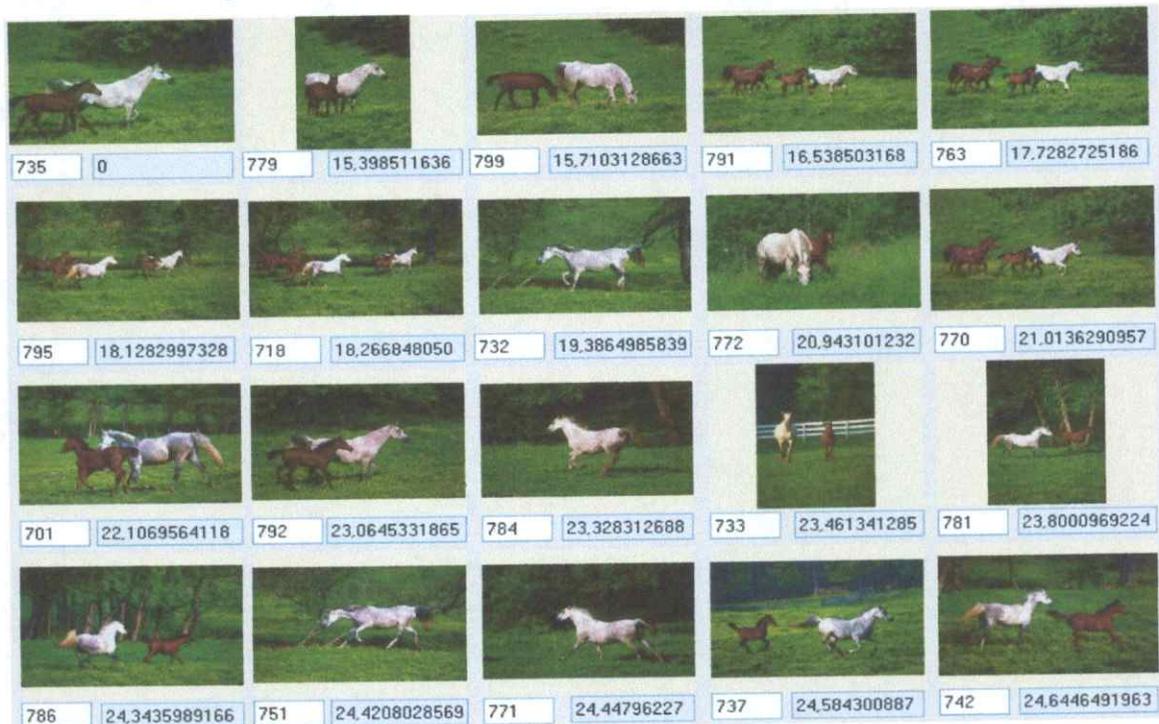
### V.3.1 Descripteur Color Layout (CLD)

Le descripteur CLD donne les meilleurs résultats pour les images simples qui contiennent des couleurs claires et simples à interpréter.

La qualité des résultats commence à diminuer avec les images de mauvaise qualité contenant des objets très compliqués, où il est difficile de distinguer et d'interpréter les couleurs. Le résultat est aussi faible pour les images monochromes, ainsi que les images d'une saturation soit très élevée ou bien très basse.

Puisque CLD prend en compte la distribution spatiale des couleurs dans l'image, un risque de pertes d'informations peut se produire, on peut par exemple trouver une image qui contient presque les mêmes couleurs que l'image de référence mais la distance de similarité entre les deux images est grande.

Prenons par exemple l'image8 de la **figure 5.1**, le résultat de la recherche en utilisant CLD est montré dans la figure suivante :



**Figure 5.3** : Résultats de la recherche en utilisant le descripteur Color Layout.

Notons aussi que le descripteur Color Layout prend en compte le degré de luminosité des couleurs dans les différentes images, car il utilise l'espace des couleurs YCbCr. Ainsi, il permet une recherche rapide dans des grandes bases d'images parce qu'il travaille avec moins de paramètres.

### V.3.2 Descripteur Scalable Color (SCD)

Contrairement au CLD, SCD ne prend pas en compte la distribution spatiale des différentes couleurs dans l'image, il donne le résultat de la recherche pour l'apparence de la couleur dans toute l'image en comparant l'histogramme des couleurs de l'image requête avec toutes les autres images.

Pour ce descripteur, la performance des résultats est presque la même que le descripteur Color Layout, sauf que ce descripteur semble plus performant pour les images contenant plusieurs objets de couleurs différentes. Ainsi, ce descripteur semble moins performant pour les images de mauvaise qualité ainsi que les images monochromes.

SCD donne des résultats bien précis dans la plupart des images car il utilise 256 paramètres, donc, il s'intéresse même aux objets de l'image.

On va montrer dans la figure suivante le résultat de la recherche des bases similaires à l'image requête en utilisant ce type de descripteur.



Figure 5.4 : Résultats de la recherche en utilisant le descripteur Scalable Color.

La recherche des images similaires pour le descripteur Scalable Color semble moins rapide par rapport au descripteur Color Layout.

### V.3.3 Descripteur Dominant Color (DCD)

Le descripteur DCD donne les meilleurs résultats dans les cas suivants :

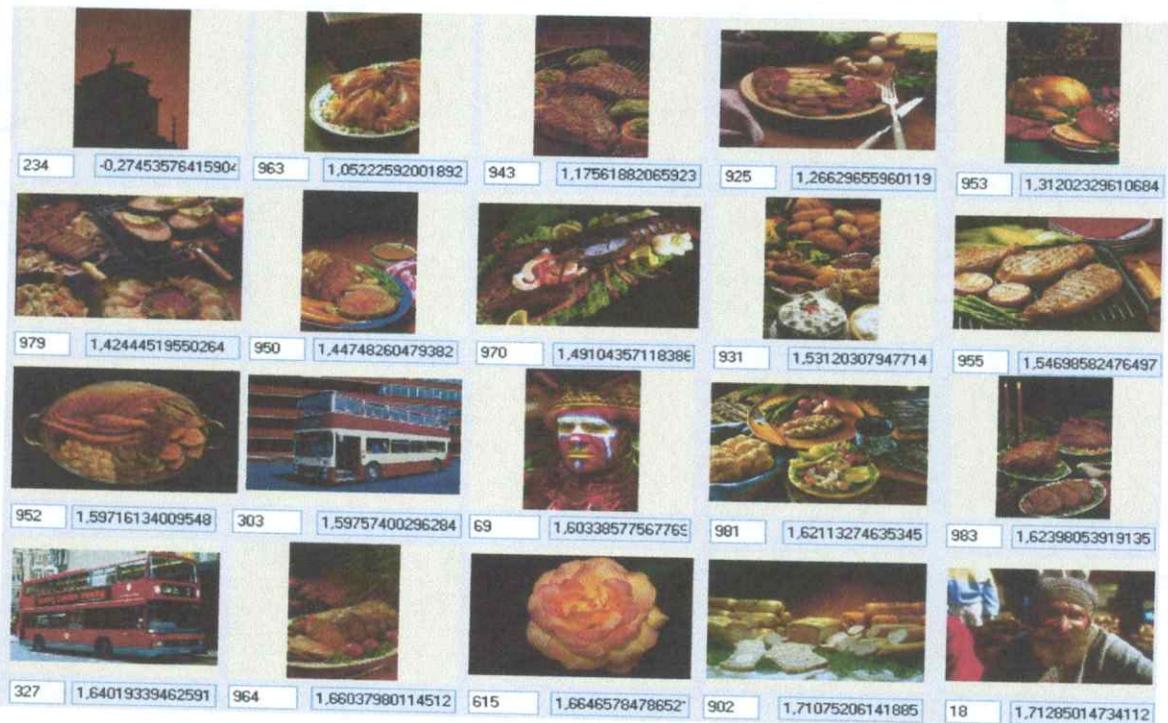
- Pour les images contenant une couleur avec un grand pourcentage. Donc, et contrairement aux descripteurs CLD et SCD, ce descripteur fonctionne bien pour les images monochromes.
- La qualité des résultats est bonne pour les images contenant moins de couleurs dominantes par rapport aux images avec plus de couleurs dominantes (la qualité diminue avec l'augmentation du nombre de couleurs dominantes dans l'image).

Par contre, DCD semble un peu mauvais pour une image qui contient plusieurs couleurs dominantes avec des pourcentages égaux.

Les résultats de la recherche obtenus pour ce descripteur sont moins précis par rapport aux deux descripteurs SCD et CLD, la raison est due au seuil de similarité entre deux couleurs dans l'image.

Si on cherche les images similaires à une image qui contient une couleur dominante avec un pourcentage important, ce descripteur trouve toutes les images contenant cette couleur, sans tenir compte des autres couleurs.

Voici un exemple des images similaires trouvés respectivement pour l'image5 qui contient 2 couleurs dominantes, et l'image1 qui contient 7 couleurs dominantes.



(a) Résultats de la recherche (image requête contient 2 couleurs dominantes)



(b) Résultats de la recherche (image requête contient 7 couleurs dominantes)

Figure 5.5: Résultats de la recherche en utilisant le descripteur Dominant Color (DCD).

V.3.4 Résultats de la combinaison

D'après les différents tests pour la recherche en utilisant la combinaison entre les descripteurs, nous avons rencontré plusieurs situations. Par exemple, pour une recherche par combinaison des deux descripteurs SCD et CLD, plusieurs cas sont possibles :

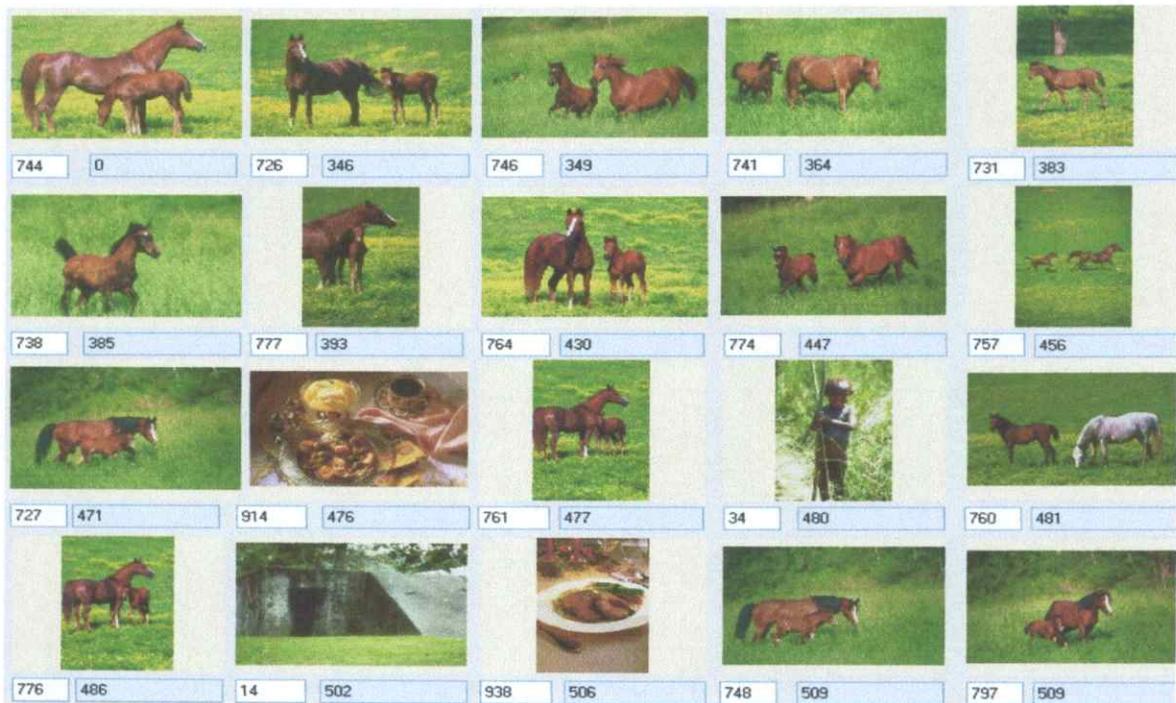
- Soit le résultat de la recherche est meilleur que la recherche en utilisant chacun uniquement, c'est le cas le plus répandu surtout pour la méthode de filtrage.
- Parfois, la combinaison donne des meilleurs résultats par rapport à un descripteur, mais ce n'est pas le cas pour l'autre. Dans ce cas, la combinaison élimine les avantages des descripteurs.

Exemple 1:

Dans l'exemple suivant, on va montrer comment la combinaison entre CLD et SCD améliore les résultats de la recherche par rapport à la recherche en utilisant chacun d'eux uniquement. Pour cette raison, on a choisi l'image 10 de la figure 5.1.



(a) résultats de la recherche en utilisant CLD.



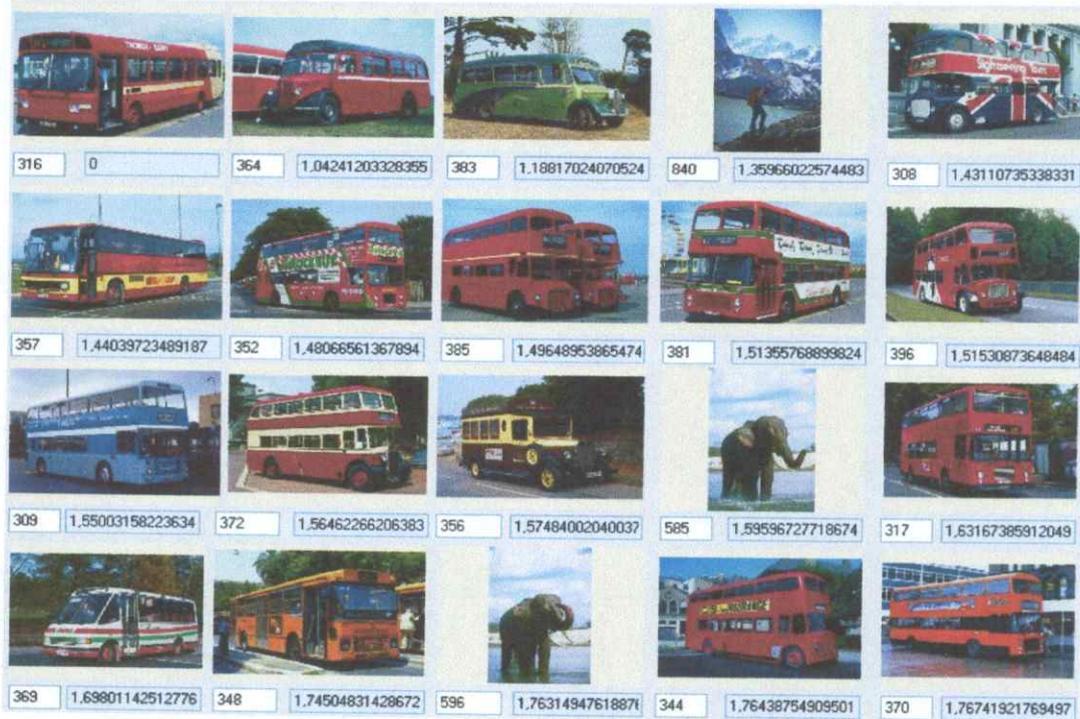
(b) résultats de la recherche en utilisant SCD.



Figure 5.6 : résultats de la recherche en combinant SCD et CLD (filtrage).

**Exemple2 :**

Dans cet exemple, on va montrer les résultats obtenus par la combinaison des deux descripteurs CLD et DCD, en utilisant la méthode de poids.



(a) résultats de la recherche en utilisant DCD.



(b) résultats de la recherche en utilisant CLD.



Figure 5.7: résultats de la recherche en combinant DCD et CLD (poids).

## V.4 ÉTUDE DES RÉSULTATS

Pour évaluer les différentes approches d'indexation et de recherche, nous avons comparé l'ensemble des résultats obtenus avec les images requêtes de la **figure 5.1**.

Nos critères de comparaison sont les suivants:

- Le nombre d'images similaires successives.
- Le nombre d'images satisfaisantes parmi les 10 premières (nb = 10).
- Le nombre d'images satisfaisantes parmi les 15 premières (nb = 15).
- Le nombre d'images satisfaisantes parmi les 20 premières (nb = 20).

Cette évaluation rentre dans la boucle de pertinence par l'utilisateur. Elle permet de qualifier les résultats obtenus pour toute la base d'images. Ainsi, ces mesures permettent le choix de descripteurs adaptés à la recherche par le contenu couleur.

### V.4.1 Descripteur Color Layout

Les résultats de la recherche obtenus pour les 10 images considérées, en utilisant le descripteur CLD sont montrés dans le tableau suivant. Les résultats obtenus diffèrent d'une image à l'autre.

Image	Nombre d'images similaires successives	Nb = 10	Nb = 15	Nb = 20
Image1	11	10	13	16
Image2	9	9	14	19
Image3	14	10	14	18
Image4	8	9	12	14
Image5	1	1	3	4
Image6	8	9	12	14
Image7	3	8	8	10
Image8	40	10	15	20
Image9	12	10	13	13
Image10	3	6	8	10

Tableau 5.1 : résultats de la recherche en utilisant Color Layout.

Pour certaines images, il est fréquent de trouver parmi les images jugées similaires des images dont le contenu est jugé différent.

### V.4.2 Descripteur Scalable Color

Les résultats sont jugés meilleurs par l'utilisateur que ceux du descripteur CLD. Dans la plupart des cas, les 20 premières images sont jugées similaires. Le tableau suivant donne les résultats obtenus pour la recherche en utilisant le descripteur Scalable Color :

Image	Nombre d'images similaires successives	Nb = 10	Nb = 15	Nb = 20
Image1	7	9	13	15
Image2	12	10	13	17
Image3	29	10	15	20
Image4	21	10	15	20
Image5	1	5	8	9
Image6	16	10	15	16
Image7	25	10	15	20
Image8	47	10	15	20
Image9	21	10	15	20
Image10	11	10	14	18

Tableau 5.2 : résultats de la recherche en utilisant Scalable Color.

#### V.4.3 Descripteur Dominant Color

Les résultats obtenus sont jugés moins bons par rapport aux deux descripteurs CLD et SCD. En effet, on retrouve parmi les 5 premières images extraites de la base, une image avec un contenu différent de celui de l'image requête.

Les mauvais résultats de ce descripteur sont dus au fait que seul une ou deux couleurs de l'image sont retenus dans le descripteur. Il en résulte un index de couleur faiblement représentatif du contenu couleur de l'image. Ne tenant pas compte de toutes les couleurs présentes dans l'image, la comparaison des descripteurs donne des résultats non satisfaisants.

Voici les résultats qu'on obtenus pour les images considérées :

Image	Nombre d'images similaires successives	Nb = 10	Nb = 15	Nb = 20
Image1	2	7	9	11
Image2	8	9	13	15
Image3	12	10	14	17
Image4	3	6	10	11
Image5	11	10	13	15
Image6	5	5	6	7
Image7	3	6	9	13
Image8	28	10	15	20
Image9	3	7	9	10
Image10	10	10	12	12

Tableau 5.3 : résultats de la recherche en utilisant Dominant Color.

#### V.4.4 Combinaison des descripteurs

Afin d'améliorer les résultats de la recherche, nous avons proposé dans notre approche la combinaison entre ces descripteurs.

Nous comparons ci après les résultats obtenus pour la recherche en utilisant chaque descripteur uniquement, et avec combinaison des descripteurs.

Le tableau ci-dessous donne le nombre d'images similaires successives trouvés pour les images considérées, en utilisant les deux méthodes de combinaison des trois descripteurs (filtrage et poids).

Les résultats obtenus avec la combinaison entre descripteurs n'améliorent pas les résultats pour toutes les recherches. Ainsi, pour l'image3 et l'image 6, le nombre d'images similaires successives est plus grand pour Scalable Color que la combinaison par filtrage.

Les résultats obtenus pour toute la base d'images, avec la combinaison fournit de meilleurs résultats lors de la recherche que lors de l'utilisation d'un descripteur.

Image	Color Layout	Scalable Color	Dominant Color	Combinaison par filtrage (sur 40)	Combinaison par poids
Image1	11	7	2	14	18
Image2	9	12	8	19	18
Image3	14	29	12	21	32
Image4	8	21	3	24	31
Image5	1	1	11	7	1
Image6	8	16	5	8	14
Image7	3	25	3	25	31
Image8	40	47	28	40	51
Image9	12	21	3	24	30
Image10	3	11	10	10	10

Tableau 5.4 : résultats de la recherche (comparaison des résultats).

## CONCLUSION

Tout au long de ce chapitre, nous avons présenté une analyse comparative des résultats obtenus pour la recherche d'images basée sur les trois descripteurs couleurs Color Layout, Scalable Color, Dominant Color, ainsi que les résultats de la recherche en utilisant la combinaison entre les descripteurs. Nous avons discuté ces résultats qui nous permettent de choisir le descripteur de recherche selon nos besoins.

Nos résultats ont en particulier mis en évidence que l'un des descripteurs est à éviter du fait de sa faible représentation du contenu couleur : c'est le descripteur dominant color. Ce dernier serait adapté si l'objectif est la recherche d'une couleur dans l'ensemble des images de la base.

La combinaison des descripteurs amène une amélioration notable des résultats. En particulier, le filtrage qui permet de faire un tri préalable de la base et de ne rechercher que dans un ensemble restreint d'images celles qui sont similaires à la requête.

*Conclusion générale et  
perspectives*



Plusieurs études traitent de la recherche d'images par leur contenu. En effet, l'image comprend une quantité importante d'information que l'homme utilise pour l'identifier et la décrire. Une description textuelle n'est donc pas toujours utile ni facile à déployer. Dans ce cadre, au lieu d'indexer les images par une description textuelle (mots), on se dirige vers sa description par son contenu. Pour ce faire, on doit au préalable extraire des indices visuels de l'image. Ce sont ces descripteurs qui permettront d'indexer l'image par son contenu.

Notre travail s'inscrit dans ces méthodes de recherche d'image par le contenu (dénommées *content-based image retrieval*), où la phase d'indexation n'implique à aucun moment l'utilisation du langage. Notre approche consiste à extraire les paramètres décrivant les couleurs présentes dans les images considérées. Les images de notre base sont identifiées par les valeurs de ses paramètres (dites signatures ou descripteurs).

Le processus de recherche s'effectue alors, non pas à partir d'une requête textuelle, mais à partir d'une image « requête » pour laquelle les paramètres de couleur sont calculés. L'image requête est alors « placée » dans l'espace de référence où sont présentes les images indexées. Les images les plus proches de l'image requête dans cet espace de référence (et ayant donc des valeurs de descripteurs similaires) sont alors présentées en résultat à l'utilisateur comme étant les images les plus proches visuellement de l'image requête.

La couleur est une caractéristique forte qu'utilise l'humain pour différencier les objets. Ainsi, c'est grâce à la couleur que l'on différencie deux objets de même forme (une table verte d'une table rouge). Plusieurs travaux ont montré que la couleur est un descripteur efficace des contenus images.

Il existe cependant de nombreuses approches pour mesurer la couleur dans une image. Laquelle de ces mesures est la plus appropriée pour décrire de manière pertinente une image ? C'est pour répondre à cette question que ce travail d'étude a été initié par le laboratoire LRIA de l'USTHB.

Dans notre travail, nous avons élaboré un système d'extraction des descripteurs couleur selon la norme MPEG-7. Nous avons étudiés trois descripteurs : descripteur Color Layout (CLD), Scalable Color (SCD), et Dominant Color (DCD).

Notre étude a pour objectif de « rendre compte » de la perception humaine de la couleur. Nous proposons donc pour la recherche d'images par contenu différentes approches : un descripteur seul, ou une combinaison pondérée des descripteurs, ou un pré filtrage par un descripteur avant de réaliser la recherche par le second. L'objectif est de permettre la mise en évidence des paramètres couleur qui se rapprochent le plus de la perception humaine. C'est l'objectif principal que c'est assigné le projet du laboratoire LRIA.

Pour y parvenir, nous avons développé notre système avec différentes combinaisons de descripteurs. Puis nous avons mené une série de tests avec une base d'images (benchmark Corell11). Les divers résultats obtenus montrent que combiner les descripteurs permet d'obtenir une meilleure recherche par le contenu.

Les différents résultats obtenus montrent que le descripteur DCD, contrairement aux descripteurs CLD et SCD, n'est adapté que pour les images monochromes (une couleur dominante).

Les résultats des descripteurs CLD et SCD se rapprochent. SCD reste plus performant pour les images contenant une grande variété d'objets de couleurs différentes. Il est cependant plus sensible à la mauvaise des images (bruits).

La combinaison des descripteurs, en utilisant la méthode de poids, améliore sensiblement les résultats. Cependant, le filtrage donne des résultats plus satisfaisants.

Les diverses évaluations ont été menés avec l'utilisateur humain. Il a permis de détecter les images pertinentes de celles jugées inadéquates. Il sera donc nécessaire dans les travaux suivants d'introduire la boucle de pertinence dans l'évaluation des résultats de la recherche. En particulier, pour se détacher de la subjectivité de chaque utilisateur, une multiplicité de mesures et une pondération des résultats obtenus permettrait de disposer d'une mesure de la qualité de la recherche.

Par ailleurs, les caractéristiques couleurs que la norme MPEG-7 considère sont insuffisantes pour décrire la richesse d'une telle caractéristique. Il serait intéressant d'intégrer d'autres caractéristiques de couleur. Une étude préalable des descripteurs couleurs existants sera nécessaire.

# *Bibliographie*

---

# Bibliographie

---

- [1] «Généralités sur le traitement d'images»  
<http://www.kaddour.com/chap1/chap1.htm>
- [2] «Analyse Multimédia»  
Sahbi SIDHOM, MCF. Université Nancy 2  
Email : [Sahbi.Sidhom@loria.fr](mailto:Sahbi.Sidhom@loria.fr)  
(Document Powerpoint Internet)
- [3] «Infographie» (document internet)  
<http://www.commentcamarche.net/contents/video/affich.php3>
- [4] Thèse : «Indexation d'image par le contenu couleur par la méthode d'histogramme».  
Réalisé par : Mr Ahcene Belhocine & Mr Rih Hamdane, Université de Boumerdes,  
Septembre 2006
- [5] [http://perso.univ-lr.fr/pcourtel/espardon/site\\_web/](http://perso.univ-lr.fr/pcourtel/espardon/site_web/)  
(Document Internet)
- [6] document internet  
<http://www.inrp.fr/lamap/scientifique/optique/essentiel/couleurs.htm>
- [7] «Segmentation d'images a base Topologique»  
Luc Brun, thèse de doctorat, université de Bordeaux, 1996
- [8] Thèse : «Indexation et recherche d'image sous différents espaces couleurs».  
Réalisé par : Mlle MEBARKI Nesrine & Mlle KHETTAR Lamia, USTHB, promotion  
2006/2007
- [9] mémoire : «Indexation et recherche d'images par le contenu couleur méthode  
d'histogramme»  
Mlle NEZELFAR Hayet & CHAMEKH Wassila, USTHB, promotion 2002/2003

# Bibliographie

---

- [10] «Développement d'une méthode d'indexation spatiale pour la recherche d'images par le contenu»  
Mémoire présenté à l'École Nationale d'Ingénieurs de Sfax par Ali KARRAY, 2006
- [11] «Indexation symbolique d'images»  
Thèse de doctorat : Stéphane Bissol, UFR 2005
- [12] «Indexation et recherche d'image par le contenu»  
Mémoire de master, Dr. Alain BOUCHER, HANOI 2005
- [13] «Comment extraire la sémantique d'une image ?»  
Alain Boucher & Thi-Lan Le  
Institut de la Francophonie pour l'Informatique, Hanoi, Vietnam  
Email : [alain.boucher@auf.org](mailto:alain.boucher@auf.org)
- [14] «indexation et recherche d'image par le contenu»  
Mr DJEFFAL Ismail & ARIB Salim, USTHB, promotion 2006/2007
- [15] «Le standard MPEG7»  
GWINNER C.Roland & LALAURETTE Sébastien  
[http:// le standard MPEG7.pdf](http://lestandardmpeg7.pdf)
- [16] «Gestion d'une bibliothèque virtuelle d'images sous Fedora».  
Diplômant : Patrick Monbaron  
Université de Genève Yverdon, le 17 décembre 2007  
[http:// PMonbaronRrtFinal.pdf](http://PMonbaronRrtFinal.pdf)
- [17] "The MPEG-7 Color Descriptors"  
Jens-Rainer Ohm (RWTH Aachen, Institute of Communications Engineering)  
[http:// vcip2003-mpeg7.pdf](http://vcip2003-mpeg7.pdf)
- [18] «Du monde de l'audiovisuel vers le monde du Web» (Doc Internet)  
[http:// websidestory.com/privacy/](http://websidestory.com/privacy/)

# Bibliographie

---

- [19] "Color and Texture Descriptors"  
Membres IEEE : B. S. Manjunath, Akio Yamada, Jens-Rainer Ohm, Vinod V. Vasudevan, Juin 2001
- [20] <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.html> (Document Internet)
- [21] [http:// pages.videotron.com/danjean/MPEG7.html](http://pages.videotron.com/danjean/MPEG7.html) (Document Internet)
- [22] "MPEG 7 Color Structure Descriptor for Visual Information Retrieval Project VizIR<sup>1</sup>"  
Adis Buturovic, Université de Vienna, décembre 2005  
[http:// COLOR STRUCTURE DESCRIPTOR.pdf](http://COLOR_STRUCTURE_DESCRIPTOR.pdf)  
Email : [e0125423@student.tuwien.ac.at](mailto:e0125423@student.tuwien.ac.at)
- [23] "Color descriptors from compressed images"  
Vinay Modi Ecole d'informatique, université d'Edinburgh, UK (Document internet)
- [24] "Introduction to MPEG-7: Multimedia Content Description Interface"  
Jim Errico & Peter van Beek, December 2001  
Email : [jerrico@sharplabs.com](mailto:jerrico@sharplabs.com)