



Faculté de Technologie

Département de Télécommunication

Mémoire

Présenté Pour l'obtention du diplôme de Master

Spécialité : Télécommunication

Option : Réseaux et télécommunications

Présenté par :

KHECHAI HANANE

MERINE IKRAM

Thème :

**Mise en place d'un système de recherche
d'image par le contenu**

Soutenu le : 07 / 07 / 2021, devant le jury composé de :

Président : M^r M.BAHRI MCB U.A.B Tlemcen

Examineur : M^r M.KHELADI MCA U.A.B Tlemcen

Encadreur : M^r R.BOUABDALLAH MAA U.A.B Tlemcen

Remerciement :

Tout d'abord, je remercie infiniment le bon dieu, tout-puissant, de la bonne santé, la volonté et la patience qu'il m'a donné tout le long de mes études.

Nos remerciements à Notre encadreur Monsieur BOUABDELLAH REDA Qui nous a guidé tout au long de Notre travail, avec son expérience, son savoir-faire, sa patience, son dévouement et ces conseils précieux Durant la réalisation de ce travail.

Nous tenons à remercier Madame BELAROUCI SARA pour son aide précieuse, sa lecture attentive du manuscrit.

Nos remerciements s'adressent aussi aux membres de Jury qui nous ont fait l'honneur de juger ce modeste travail.

Nous exprimons également notre gratitude à tous les enseignants qui ont contribué à notre formation.

Enfin, je ne pourrai jamais assez remercier mon père et ma mère, que je porte toujours avec moi. Sans leur grande confiance, leur soutien et leurs prières, je n'aurais pas pu mener à bien mes projets.

Enfin, nos remerciements les plus sincères à tous ceux qui ont contribué de près ou de loin à la concrétisation de ce mémoire.

Dédicace :

Je tiens c'est avec grande plaisir que je dédie ce modeste travail :

A mes très chers parents pour leur soutien, leur patience, leur encouragement durant mon parcours scolaire, que Dieu les protège.

A mes chers frères Nabil et Amine.

A ma grande sœur Nacira et son mari, et ses petites filles Rahaf Assil et Loujaine Cirine.

A toute la famille KHECHAI et BENMANSOUR, A mes amies proches Khaoula, Chahra, Imane, Ibtissem.

A mon binôme Ikram avec laquelle j'ai partagé cette expérience.

A tous ceux que j'aime, et tous ceux qui m'aiment.

Sans oublier toute la promotion de Master RT 2020/2021.

KH. Hanane

Dédicace :

Je dédie du fond du cœur ce mémoire :

A mes chers parents et à ma grand-mère, à qui je dois tout, pour leur amour, leurs prières et leur soutien qui m'ont été donnés tout au long de mes études.

A ma sœur chérie et mes frères, je la dédie aussi à mon cher époux, qui a partagé avec moi en tous sens ces précieuses années, je le remercie pour ses encouragements, son soutien, et pour toute sa compréhension et sa grande compagnie.

A ma chérie Amel, ma meilleur amie.

M. Ikram

Résumé

Avec l'expansion de l'informatique et du multimédia, une problématique nouvelle est apparue c'est de gérer les quantités énormes et croissantes de données (plus particulièrement les bases d'images), la structuration, le stockage et la recherche sont les problèmes les plus rencontrés lors de la manipulation de ces bases d'images, il y a une difficulté d'associer une sémantique à une image, car les images sont nommées d'une manière générique (img001.jpg).

Cette thèse s'inscrit dans la problématique de l'indexation et la recherche d'images par le contenu dans des bases d'images volumineuses. Dans notre travail nous avons suggéré une technologie intégrée de récupération d'image basée sur le contenu qui extrait à la fois la couleur et la texture. Pour extraire la fonction de couleur, le moment de couleur (CM) est utilisé sur les images en couleur et pour extraire la fonction de texture, le motif binaire local (LBP) est effectué sur l'image en niveaux de gris, il est généralement utilisé dans la reconnaissance faciale mais nous allons l'utiliser pour les images naturelles. Ensuite ces deux descripteurs sont combinés pour former un seul vecteur descripteur. En fin de compte on compare le vecteur descripteur des images de base de données avec les images de requête la mesure de similarité qui est la distance euclidienne. Ces approches combinées fournissent des systèmes de récupération efficace et moins complexe.

Mots clés : traitement d'image, recherche d'image par le contenu(CBIR), l'indexation,

Descripteurs d'images, mesures de similarité.

Abstract

With the expansion of computing and multimedia, a new problem has emerged is to manage the huge and increasing quantities of data (more particularly image databases), structuring, storage and research are the most encountered problems when handling these image databases, there is a difficulty of associating semantics with an image, because the images are named generically (img001.jpg).

This thesis is part of the problem of indexing and searching images by content in large image databases. In our work we suggested an integrated content-based image recovery technology that extracts both color and texture. To extract the color function, the color moment (CM) is used on color images and to extract the texture function, the local binary pattern (LBP) is performed on the grayscale image, it is usually used in facial recognition but we will use it for natural images. Then these two descriptors are combined to form a single descriptor vector. Ultimately we compare the descriptor vector of the database images with the query images the similarity measure which is the Euclidean distance. These combined approaches provide efficient and less complex recovery systems.

Keywords : image processing, Content-Based Image Retrieval (CBIR), indexing, Image

descriptors, similarity measures.

Table des matières

- I. Remerciement**
- II. Dédicaces**
- III. Résumé**

Introduction générale1

CHAPITRE 01

Généralités sur le traitement d'images

1 . Introduction3

2 . Notions de base3

 2.1 Définition de l'image3

 2.2 Les types d'images3

 2.2.1 L'image en pixels (Bitmap).....3

 2.2.2 L'image vectorielle :.....4

 2.3 Formats d'images4

 2.3.1 JPEG (Joint Fotographic Expert Group).....4

 2.3.2 PNG (Portable Network Graphic)4

 2.3.3 AI (Adob Illustrator).....4

 2.3.4 PDF (Portable Document Format).....5

 2.3.5 BMP5

 2.3.6 GIF (Graphic Interface Format)5

 2.3.7 TIFF (Tag Image File Format)5

 2.3.8 EPS (Encapsulated Post Script).....5

3. Caractéristiques d'une image5

 3.1 Dimension.....5

 3.2 Pixel.....5

 3.3 Texture.....5

 3.4 Résolution.....6

 3.5 Bruit.....6

 3.6 Histogramme.....6

 3.7 Luminance6

 3.8 Contraste.....6

3.9 Niveaux de gris	6
3.10 Images en couleurs	7
4 .Pré-traitement des images	7
4.1 La modification d’histogramme.....	7
4.1.1 Expansion de dynamique	7
4.1.2 Egalisation d’histogramme.....	8
4.2 L’augmentation de contraste.....	9
4.3 La suppression du bruit	9
5. Filtrage	10
5.1 Filtres linéaires.....	10
5.1.1 Filtre moyenneur (mean filtering ou box filtering).....	10
5.1.2 Filtre gaussien (gaussian filtering).....	11
5.2 Filtres non linéaires	11
5.2.1 Filtre médian (median filter)	11
5.2.2 Filtres morphologiques	12
6 . Segmentation.....	12
6.1 La définition.....	12
6.2 Définition formelle de la segmentation	13
6.3 Les méthodes de segmentation	13
6.3.1 Segmentation par régions	14
6.3.1.1 croissance de région	14
6.3.1.2 Segmentation par fusion de régions	15
6.3.1.3 Segmentation par division de régions	15
6.3.1.4 Segmentation par division –fusion.....	15
6.3.2 Segmentation par contours.....	16
6.3.2.1 Méthodes dérivatives.....	16
6.3.2.2 Méthodes analytiques	19
6.3.2.3 Méthodes déformables.....	19
6.3.2.4 Les limites de segmentation par contour	20
7 . Conclusion.....	20

CHAPITRE 02

Systèmes de recherche d'image par le Contenu (CBIR)

1. Introduction	21
2 . Composants d'un CBIR	22
2.1 Les bases d'images utilisées	22
2.1.1 La base de Wang.....	22
2.1.2 COIL (Columbia Object Image Library)	23
2.1.3 Pollen	24
2.1.4 CURET(Columbia Utrecht Reflectance and Texture Database)	24
2.1.5 La base de FeiFei.....	25
2.2 L'indexation.....	25
2.3 La gestion des index	26
2.4 Les requêtes	26
2.4.1 Requête par mots clés	26
2.4.2 Requête par esquisse.....	27
2.4.3 Requête par le contenu	27
2.4.4 Requête par l'exemple	27
2.5 Analyse de la requête	28
2.6 Mise en correspondance requête / base	28
2.7 La présentation des résultats	28
3. Mesures pour évaluer un système.....	28
3.1 Rappel et précision (en anglais : Recall and Precision)	29
3.2 La courbe de rappel et précision	30
4. Conclusion	31

CHAPITRE 03

Descripteurs d'image et mesures de similarité

1. Introduction	32
2. Descripteurs d'images.....	32
2.1 Descripteurs de couleur	32
2.1.1 Les espaces de couleurs	32

2.1.1.1 Espace RVB ou RGB en anglais	32
2.1.1.2 Espace TSV.....	33
2.1.1.3 Système XYZ.....	34
2.1.1.4 Système $L*u*v$	34
2.1.1.5 Système $L*a*b$	35
2.1.2 L'Histogramme.....	35
2.1.3 Les moments statistiques	36
2.1.4 Corrélogramme de couleurs	36
2.1.5 Cohérence spatial.....	37
2.1.6 Descripteur par couleurs dominantes.....	38
2.2 Les descripteurs de textures.....	38
2.2.1 Les méthodes statistiques	39
2.2.1.1 Les matrices de longueur de plages (MLDP)	39
2.2.1.2 Matrice de cooccurrence.....	40
2.2.1.3 La méthode de différence de niveaux de gris (GLDM)	41
2.2.1.4 Caractéristique de Tamura.....	41
2.2.2 Les méthodes fréquentielles	41
2.2.2.1 Transformée de Fourier	41
2.2.2.2 Transformée en ondelettes.....	41
2.2.2.3 Filtres spatiaux (Spatial domain filters)	42
2.2.2.4 Filtres de Gabor.....	42
2.3 Les descripteurs de formes	43
3. Mesure de similarité	46
3.1 Les méthodes de calcul	46
3.1.1 Distance de Minkowski.....	46
3.1.2 Distance quadrature	47
3.1.3 Distance cosinus	47
3.1.4 Distance de Mahalanobis	47
3.1.5 Distance de bhattacharya	48
3.1.6 Distance de Kullback Leiber(KL)	48
3.1.7 Distance de Kolmogorov Smirnov	48
3.1.8 Distance de Cramer Von Mises	48
3.1.9 Intersection d'histogramme.....	48

3.1.10 Earth Mover Distance(EMD)	49
3.1.11 Divergence de Jeffrey (JD)	49
4. Conclusion	49

CHAPITRE 04

Travail effectué

1. Introduction	50
2. Méthode utilisée	50
3. Architecture.....	50
3.1 Le prétraitement	51
3.2 L'extraction de descripteur	51
local binary pattern (LBP)	51
les moments de couleur	53
4. Mesure de similarité	53
5. L'algorithme.....	54
6. Bases d'images utilisé.....	55
7. Résultat expérimental	55
8. Conclusion	58
Conclusion générale	59

Bibliographie

Liste des figures

Chapitre 01

Généralités sur le traitement d'images

Figure 1.1 : Exemple d'image en pixels.....	03
Figure 1.2 : Exemple d'image vectorielle.....	04
Figure 1.3 : Expansion de dynamique	08
Figure 1.4 : Egalisation d'histogramme.....	09
Figure 1.5 : Application du filtre moyen.	10
Figure 1.6 : Application du filtre gaussien.....	11
Figure 1.7 : Principe du filtre médian.....	12
Figure 1.8 : Exemple de segmentation d'image.....	13
Figure 1.9 : Approche région et approche contour.....	14
Figure 1.10 : Segmentation par croissance de régions	14
Figure 1.11 : les modèles de contours	16
Figure 1.12 : contour et ses dérivées.....	17
Figure 1.13 : Détection de contour.....	18
Figure 1.14 : contour détecté par Laplacien.....	19

Chapitre 02

Systèmes de recherche d'image par le Contenu (CBIR)

Figure 2.1 : Principaux composants d'un Système de Recherche par le Contenu.....	22
Figure 2.2 : 10 classes de la base de Wang (Deselaers, 2003)	23
Figure 2.3 : Les objets utilisés dans COIL-100(Deselaers, 2003)	23
Figure 2.4 : Les objets utilisés dans COIL-20 (Deselaers, 2003)	24
Figure 2.5 : Quelques images exemples de la base de Pollens.....	24
Figure 2.6 : Quelques images exemples dans la base de CURET.....	25
Figure 2.7 : Quelques images exemples dans la base de Fei-Fei.....	25
Figure 2.8 : Un exemple de recherche d'images dans Google.....	27
Figure 2.9 : Le rappel et la précision pour une requête (Yates, 1999)	29
Figure 2.10 : Exemple de calcul rappel et précision.....	30

Chapitre 03

Descripteurs d'image et mesures de similarité

Figure 3.1 : Images couleurs dans l'espace RGB.....	33
Figure 3.2 : Représentation conique de tout l'espace TSV en un seul objet...	33

Figure 3.3 : Représentation de l'espace couleur XYZ.....	34
Figure 3.4 : Espace de couleur Lab	35
Figure 3.5 : Image couleur et son histogramme des trois composantes RVB	36
Figure 3.6 : Représentation d'un filtre de Gab.....	42

Chapitre 04 **Travail effectué**

Figure 4.1 : Architecture de travail effectué.....	50
Figure 4.2 : Descripteur LBP.....	51
Figure 4.3 : Image requête.....	55
Figure 4.4 : Résultat de moment de couleur.....	56
Figure 4.5 : Résultat de LBP.....	56
Figure 4.6 : Résultat de vecteur descripteur.....	57

Liste des tableaux

Tableau 1 : Sous-bloc de l'image de taille 3X3.....	52
Tableau 2 : Position du bit de sous-bloc de l'image.....	52
Tableau 3 : Sous-bloc 3X3 montrant la valeur binaire.....	52
Tableau 4 : Sous-bloc 3X3 valeur pondérée de la position du bit.....	52
Tableau 5 : Sous-bloc 3X3 de la valeur du descripteur LBP.....	53

Les abréviations

AI : Adob Illustrator.

CBIR : Content Based Image Retrieval.

CIE : Comission Internationale de l'Eclairage.

COIL : Columbia Object Image Library.

CURET : Columbia Ulrecht Reflectance and Texture Database.

DPI : Dot Per Inch.

EMD : Earth Mover Distance.

EPS : Encapsulated Post Script.

GIF : Graphic Interface Format.

GLDM : Gray Level Différence Method.

HSP : Hue-Saturation-Brightness.

HSV : Hue-Saturatuin-Value.

JD : Divergence de Jeffrey.

JPEG : Joint Photographic Expert Group.

KL : Kullback Leiber.

MLDP : Matrices de Longueur De Plages.

PDF : Portable Document Format.

PNG : Portable Network Graphic.

PPP : Pixel Par Pouce.

RGB : Red-Green-Blue.

RVB : Rouge-Vert-Bleu.

SGLDM : Spatial Gray Level Dépendance Method.

TIFF : Tag Image File Format.

TSV : Teinte-Satuation-Valeur.

Introduction générale

Le développement des nouvelles technologies de l'information et de la communication et l'essor du web ont entraîné un accroissement très important de l'information en terme de diversité et d'hétérogénéité. la recherche de cette information est devenue le cœur de tous les secteurs d'activités : la médecine, l'astronomie, l'histoire, la géographie le domaine de l'information et de l'audiovisuel, la surveillance, l'archivage (patrimoine culturel, musées, . . .) l'agences photographique . . . etc. Ces progrès se sont accompagnés d'une baisse des coûts des équipements informatiques qui a facilité la diffusion et l'échange de données multimédia numérisées vers le grand public. Cette masse de donnée n'aurait aucun intérêt si l'on ne pouvait pas facilement retrouver les informations recherchées. Cela a suscité un besoin en développement de techniques de recherche d'information multimédia, et en particulier de recherche d'images.

la recherche de l'information est le processus de trouver un objet de nature non structurée qui satisfait un besoin d'information au sein de collections de données de tailles très importantes stockées sur des supports de stockage. Ce domaine a évolué pour donner de nouvelles approches pour la recherche d'information de nature variée (l'information recherchée peut être un document texte, une image, une vidéo etc.).

Sur le web, les moteurs de recherches (par ex. Google) sont les principaux outils de recherche d'information. Ils scrutent le web et maintiennent une base d'index sur des documents de nature variée. Les documents image sont parmi les documents les plus recherchés sur le web. Nous nous intéressons dans notre travail à la recherche d'images. Le domaine de l'image numérique est un domaine en pleine expansion. Depuis quelques années, avec l'explosion d'Internet et aussi le développement à grande échelle de la photographie numérique, on trouve des bases d'images numériques contenant plusieurs milliers et même plusieurs dizaines de milliers d'images. Ils sont générés et stockés sur des supports de stockage et cette accumulation d'images numériques pose rapidement le problème d'indexation et de recherche d'images qui consiste à établir une correspondance entre l'image disponible et celle recherchée par l'utilisateur. Pour gérer et utiliser efficacement ces bases d'images, un système d'indexation et de recherche d'images est nécessaire.

L'indexation et la recherche d'images par le contenu est une piste prometteuse. Elle offre la possibilité aux utilisateurs d'accéder, d'interroger et d'exploiter directement les bases d'images en utilisant leur contenu ; ceci explique l'activité de recherche consacrée à ce domaine. L'idée de faciliter l'accès à des données n'est pas neuve, des techniques de recherche d'images ayant été développées à cet effet depuis la fin des années 70. Au début les systèmes de recherche et d'indexation d'images, sont basés sur l'indexation textuelle manuelle à l'aide des mots clés (l'indexation par le contexte), parmi ces approches on trouve la technique de recherche d'images à base de texte connue sous le nom « Text-based Image Retrieval » ou TBIR qui est l'approche la plus ancienne utilisée jusqu'à nos jours, elle pose sur l'annotation manuelle de chaque image par un ensemble de mots-clés décrivant leur contenu, puis d'utiliser un système de gestion de base de données pour gérer ces images.

Malgré le grand succès de cette approche pour la recherche des documents, l'annotation textuelle manuelle à l'aide des mots clés (l'indexation par le contexte) est une tâche longue et

répétitive pour l'utilisateur, surtout avec les bases d'images qui deviennent de plus en plus grandes.

C'est pour cela la recherche par le contenu a été proposée dans les années 90 qui consiste à rechercher l'image en se basant sur son contenu visuel comme le texture, la forme, la couleur.. etc. Elle s'est révélée efficace et très utile dans de nombreux domaines d'application. La recherche d'images par le contenu visuel (Content Based Image Retrieval CBIR ou RIC en français) vise à extraire directement l'information à partir de l'image, cette nouvelle modalité a ouvert des possibilités pour les utilisateurs. Ces systèmes proposent de formuler une requête au moyen d'une image exemple, en caractérisant ses propriétés visuelles de bas niveau telle que la couleur, la texture et la forme, qui sont calculées de manière automatique, et ensuite ils exploitées par le système pour comparer et retrouver des images.

Dans notre cas une image requête sera fourni par l'utilisateur pour laquelle on souhaiterait obtenir des images visuellement similaires ce qui est proposé actuellement par les moteurs de recherche tels que Google.

Organisation du mémoire

Ce mémoire est structuré en quatre chapitres :

- *Le premier chapitre*, présente une généralité de traitement d'image, il décrit la définition d'une image, leurs caractéristiques, le principe du système de traitement d'image, filtrage et quelques applications concrets sur le traitement d'image.
- *Le deuxième chapitre*, concerne la définition du domaine, il s'agit de la recherche d'images par le contenu, il explique le principe de fonctionnement des systèmes de recherche d'images par le contenu, les bases d'images utilisées et les différents requêtes.
- *Le troisième chapitre*, décrits les descripteurs d'images utilisés, tel que les espaces de couleurs, l'histogramme et les moments de Zernike. Et les différentes distances utilisé pour mesurer la similarité entres ces descripteurs.
- *Le dernier chapitre*, présente les différentes étapes de notre travail en décrivant les protocoles d'évaluation de nos approches et les résultats que nous avons obtenu.

CHAPITRE 01 :

Généralités sur le traitement d'images

1. Introduction

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

Dans ce chapitre nous présentons quelques notions de base du domaine de traitement d'image tels que : la définition d'image, les types d'image, les formats d'images, les caractéristiques d'images, pré-traitement d'images, filtrage, et en fin segmentation.

2. Notions de base

2.1 Définition de l'image

Une image est une représentation du monde. En traitement d'image, on considère qu'il s'agit d'une fonction mathématique de $R \times R$ dans R où le couplet d'entrée est considéré comme une position spatiale, le singleton de sortie comme l'intensité (couleur ou niveaux de gris) du phénomène physique. Il arrive cependant que l'image soit dite "3D" donc la fonction est de $R \times R \times R$ dans R . Les images couleurs peuvent être représentées soit par trois images représentant les trois couleurs fondamentales, soit par une image de $R \times R$ dans $R \times R \times R$.

L'image numérique est l'image dont la surface est divisée en éléments de tailles fixes appelés cellules ou pixels, ayant chacun comme caractéristique un niveau de gris ou de couleurs prélevé à l'emplacement correspondant dans l'image réelle, ou calculé à partir d'une description interne de la scène à représenter.

2.2 Les types d'images

On distingue deux types d'images :

2.2.1 L'image en pixels (Bitmap)

L'image en pixels (Bitmap) est représentée sous la forme d'une grille de deux l'axe X et Y est constitué de nombreux points: appelés pixels.

Chaque pixel trouvera un échantillon de couleur.

La définition d'une image matricielle dépend ainsi du nombre de pixels qui la composent. Plus ils sont nombreux, plus l'image sera nette : c'est la résolution (Figure 1.1)

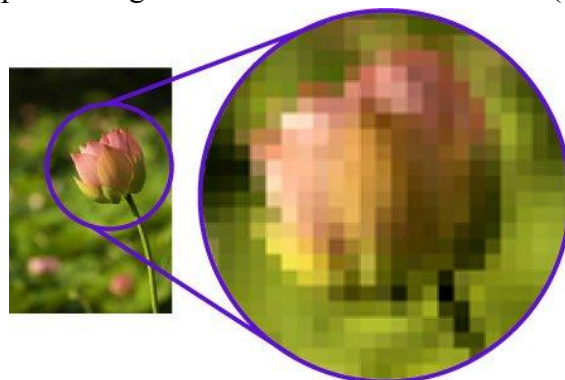


Figure 1.1 : Exemple d'image en pixels

2.2.2 L'image vectorielle :

L'image vectorielle est représentée par des formules géométriques, c'est –à-dire un ensemble de points reliés par des segments de droite, polygones, arcs de cercle, et des courbes formant un tracé.

L'avantage de ce type d'image est la possibilité d'agrandir une image vectorielle sans perdre la qualité initiale, ainsi qu'un faible encombrement. Ce type d'images est aussi utilisées sur internet pour la création de bannières publicitaires, les animations Flash, voir des sites web complet (Figure 1.2).

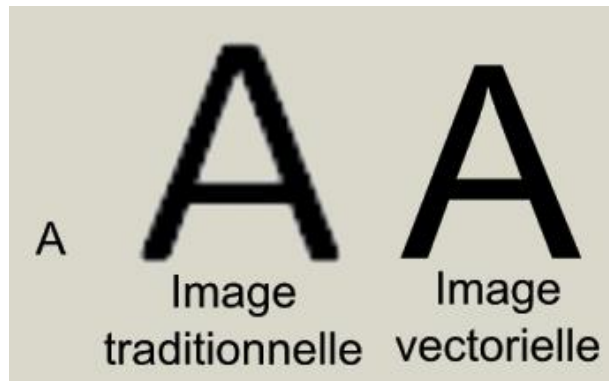


Figure 1.2 : Exemple d'image vectorielle

2.3 Formats d'images

Un format d'image est une représentation informatique de l'image, associée à des informations sur la façon dont l'image est codée et fournissant éventuellement des indications sur la manière de la décoder et de la manipuler, voici une présentation des principaux formats :

2.3.1 JPEG (Joint photographic Expert Group)

Le format JPEG est un des formats d'images les plus répandus : très utilisé pour les photographies sur Internet ou en sortie d'appareils photographiques numériques. Il concerne les images couleurs représentées sur 24bits, son principe consiste à coder l'image par blocs de 8*8 pixels. Les images sont compressées avec perte d'informations ou non, selon un taux de compression variable. [1]

2.3.2 PNG (Portable Network Graphic)

Le format PNG est le "troisième" format Internet. Beaucoup moins utilisé que le JPEG et le GIF, il présente en théorie des performances intéressantes, mais qui restent à confirmer dans la pratique quotidienne. Cette format peut intégrer des images RVB, en niveaux de gris, 256couleurs indexées et bitmap, il utilise une compression sans perte. [2]

2.3.3 AI (Adob Illustrator)

Est un format de fichier développé par Adobe pour représenter un dessin vectoriel soit en EPS ou en PDF.

2.3.4 PDF (Portable Document Format)

PDF est un format de fichier universel qui peut conserver les polices, les images, la mise en page et les graphiques. Les fichiers PDF peuvent être partagés, visualisés et imprimés par n'importe qui avec le logiciel Adobe Reader. [3]

2.3.5 BMP

Créé par Microsoft, le format BMP, en fait, le format bitmap interne du système d'exploitation Windows. Un fichier BMP peut contenir une image RVB, en couleurs indexées, en niveaux de gris ou en noir et blanc (mais convertie en 256 niveaux de gris). [2]

2.3.6 GIF (Graphic Interface Format)

Le GIF est un format comme le JPEG, qui est utilisé sur internet, il est reconnu par tous les navigateurs HTML, il présente des caractéristiques très intéressantes comme la transparence, l'affichage progressif et la possibilité de créer des animations. Un fichier GIF contient une image en 2, 4, 8, 16, 32, 64, 128 ou 256 couleurs, ces couleurs sont stockées dans une palette. [2]

2.3.7 TIFF (Tag Image File Format)

C'est certainement le format le plus connu, ce format a été créé par Microsoft et Aldus. Un fichier TIFF peut compresser des images selon les méthodes LZW, Huffman, Pckbits. Il est capable de stocker des couches alpha. [2]

2.3.8 EPS (Encapsulated Post Script)

EPS s'agit d'un format metafile créé par Adobe. Il accepte les couleurs en RVB et CMJN, et il intègre une prévisualisation de l'image. [2]

3. Caractéristiques d'une image

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

3.1 Dimension

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

3.2 Pixel

Le pixel est le plus petit point de l'image, c'est une valeur numérique représentative des intensités lumineuses. L'information présentée par le pixel est le niveau de gris (ou la couleur) prélevée à l'emplacement correspondant dans l'image réelle.

3.3 Texture

Une texture est un champ de l'image qui apparaît comme un domaine cohérent et homogène, c'est -à-dire formant un tout pour un observateur. C'est cette propriété de cohérence par l'œil humain qui sera recherchée le plus souvent par le traicteur des images, dans le but d'isoler les textures, soit pour segmenter l'image, soit pour reconnaître des régions. [4]

3.4 Résolution

La résolution de l'image est le nombre de pixels par unités de longueur. Elle s'exprime en ppp (pixel par pouce) ou dpi (dot per inch). Un pouce mesure 2.54cm. La résolution permet ainsi d'établir le rapport entre la définition en pixels d'une image et la dimension réelle de sa représentation sur un support physique (affichage écran, impression papier...). [5]

3.5 Bruit

Un bruit dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur. [6]

3.6 Histogramme

L'histogramme des niveaux de gris ou des couleurs d'une image est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleur) dans l'image. Il permet de donner un grand nombre d'information sur la distribution des niveaux de gris (couleur) et de voir entre quelles bornes est répartie la majorité des niveaux de gris (couleur) dans le cas d'une image trop claire ou d'une image trop foncée. Il peut être utilisé pour améliorer la qualité d'une image (Rehaussement d'image) en introduisant quelques modifications, pour pouvoir extraire les informations utiles de celle-ci. [7]

3.7 Luminance

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

- Des images lumineuses (brillantes).
- Un bon contraste : il faut éviter les images où la gamme de contraste tend vers le blanc ou le noir; ces images entraînent des pertes de détails dans les zones sombres ou lumineuses.
- L'absence de parasites.

3.8 Contraste

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et les régions claires de cette image. Le contraste est défini en fonction des luminances de deux zones d'images. Si L_1 et L_2 sont les degrés de luminosité respectivement de deux zones voisines A_1 et A_2 d'une image, le contraste C est défini par le rapport :

$$C = \frac{L_1 - L_2}{L_1 + L_2}$$

3.9 Niveaux de gris

Le niveau de gris est la valeur de l'intensité lumineuse en un point. La couleur du pixel peut prendre des valeurs allant du noir au blanc en passant par un nombre fini de niveaux intermédiaires. Donc pour représenter les images à niveaux de gris, on peut attribuer à chaque pixel de l'image une valeur correspondant à la quantité de lumière renvoyée. Cette valeur peut être comprise par exemple entre 0 et 255. Chaque pixel n'est donc plus représenté par un

bit, mais par un octet. Pour cela, il faut que le matériel utilisé pour afficher l'image soit capable de produire les différents niveaux de gris correspondant. Le nombre de niveaux de gris dépend du nombre de bits utilisés pour décrire la " couleur " de chaque pixel de l'image. Plus ce nombre est important, plus les niveaux possibles sont nombreux.

3.10 Images en couleurs

Même s'il est parfois utile de pouvoir représenter des images en noir et blanc, les applications multimédias utilisent le plus souvent des images en couleurs. La représentation des couleurs s'effectue de la même manière que les images monochromes avec cependant quelques particularités. En effet, il faut tout d'abord choisir un modèle de représentation. On peut représenter les couleurs à l'aide de leurs composantes primaires. Les systèmes émettant de la lumière (écrans d'ordinateurs,...) sont basés sur le principe de la synthèse additive : les couleurs sont composées d'un mélange de rouge, vert et bleu (modèle R.V.B.)

4 .Prétraitement des images

Le prétraitement d'images est l'ensemble des opérations effectuées sur une image, soit pour l'améliorer, soit pour la restaurer. D'une manière générale, les méthodes de prétraitement visent soit à renforcer la ressemblance des pixels d'une même région, soit au contraire à accentuer les différences de pixels provenant de régions différentes. Les techniques utilisées apparaissent dans le traitement des images pour des applications scientifiques, plutôt que pour les applications de divertissement ou de décoration. Alors on va pratiquer et comparer différentes méthodes de prétraitement ou d'amélioration d'images. Ces prétraitements consistent d'une part à modifier l'apparence d'une image pour qu'un observateur puisse plus facilement en extraire des informations, et d'autre part à supprimer l'information inutile (le bruit) des images pour renforcer l'information utile en vue d'un traitement ultérieur.

Les techniques de prétraitements les plus courantes qu'on va présenter sont :

- La modification d'histogramme.
- L'augmentation de contraste.
- La suppression du bruit.

4.1 La modification d'histogramme

On cherche à améliorer l'image en lui appliquant une transformation ponctuelle d'intensité A tout pixel d'intensité s on associe une intensité $t = T(s)$ (as). Du fait de leur caractère ponctuel (les pixels sont traités individuellement), les méthodes de transformation d'histogramme n'affectent pas la forme des régions mais modifient uniquement l'apparence visuelle de l'image.

Dans ce qui suit, on va présenter deux types de modification d'histogramme :

- Expansion de dynamique.
- Egalisation d'histogramme.

4.1.1 Expansion de dynamique

Cette méthode consiste à utiliser au mieux l'échelle des niveaux de gris disponible sur le système d'acquisition d'images.

Soit A [i, j] l'image de départ, A' [i, j] image après transformation et [a₀, a₁] l'intervalle des intensités présentes dans l'image et soit [a_{min}, a_{max}] l'intervalle disponible.

L'expansion du dynamique correspond à la transformation T : a's=α+β as telle que :

$$\forall a \in [a_0, a_1] \text{ as} \rightarrow a's \in [a_{min}, a_{max}]$$

$$\alpha = \frac{a_{min} * a_0 - a_{max} * a_1}{a_1 - a_0} ; \beta = \frac{a_{max} - a_{min}}{a_1 - a_0}$$

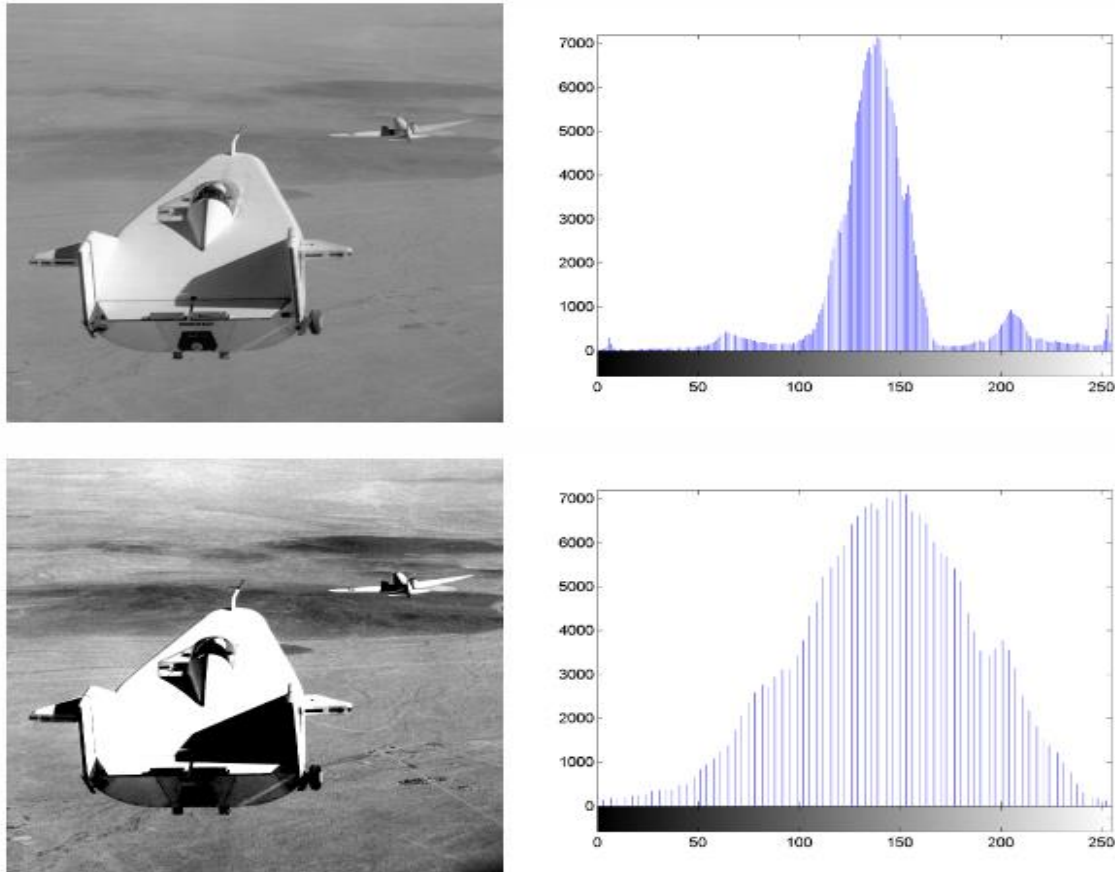


Figure 1.3 : Expansion de dynamique

4.1.2 Egalisation d'histogramme

Cette transformation consiste à rendre le plus plat possible l'histogramme de niveaux de gris de l'image. L'image A[i, j] est considérée comme un ensemble de réalisation d'une variable aléatoire A, admettant une densité de probabilité f(a) et une fonction de répartition F(a), f représente l'histogramme normalisé de A[i, j]. On cherche une transformation T continue dérivable par morceaux et strictement croissante telle que la variable aléatoire B=T(A) soit uniformément répartie dans l'intervalle [b_{min}, b_{max}]. Soit g(b) la densité de probabilité de B, et T' la transformation dérivée de T :

$$g(b) = \begin{cases} f(a) \frac{1}{T'(a)} = \frac{1}{b_{max} - b_{min}} & \text{si } b_{min} < b < b_{max} \\ 0 & \text{ailleurs} \end{cases}$$

$$T'(a) = (b_{max} - b_{min}) \cdot f(a) \quad , b = T(a) \quad , a \in [a_{min}, a_{max}]$$

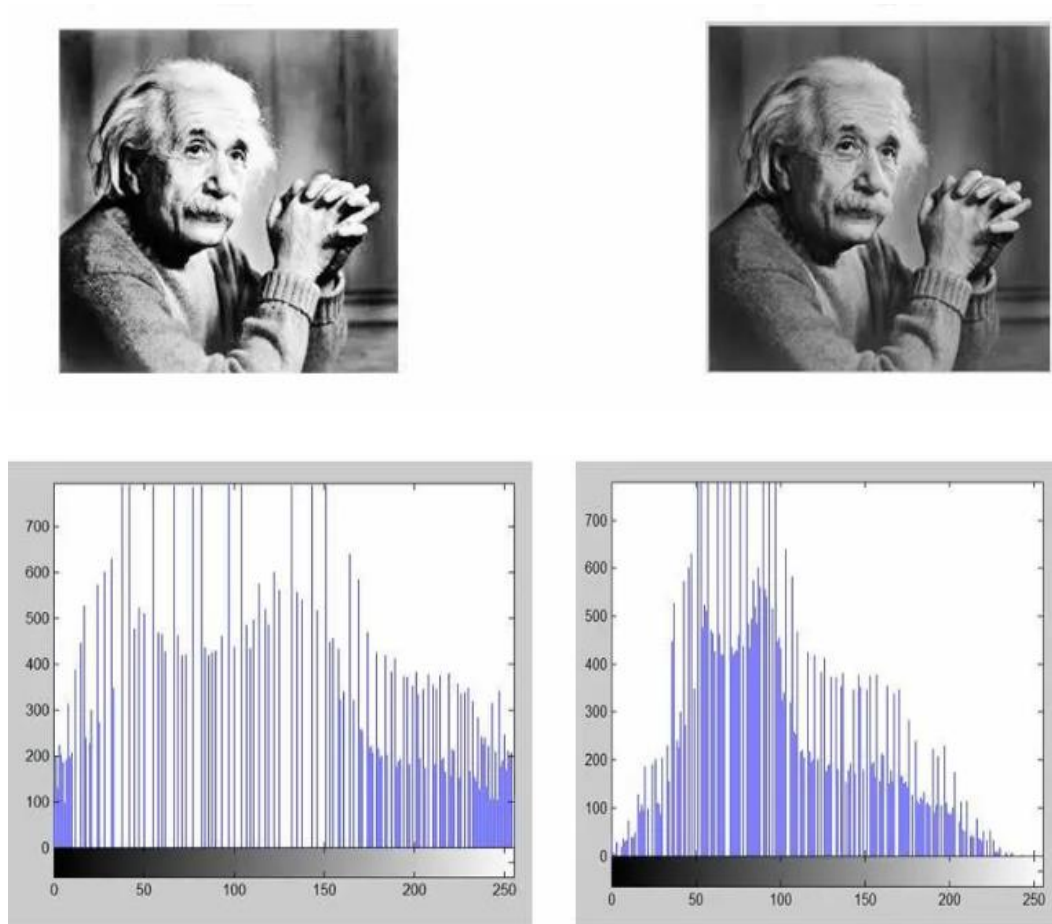


Figure 1.4 : Egalisation d'histogramme

4.2 L'augmentation de contraste

Une famille de traitements a pour objectif de donner à l'image un plus grand contraste. Ces méthodes procèdent de trois façons différentes :

- Par étirement d'histogramme : on cherche à exploiter au mieux la dynamique de l'histogramme, tout d'abord en utilisant toute l'échelle de gris ou de couleurs disponible, mais aussi en modifiant l'histogramme de façon à le transformer en un histogramme de référence. Les plus utilisés sont les histogrammes plats et hyperboliques ;
- Par filtrage passe-haut de l'image : l'idée de base est ici de réduire l'importance du terme continu et des basses fréquences. Un point délicat de ces méthodes est de définir le gabarit du filtre utilisé ;
- Par des méthodes locales : ces méthodes modifient localement l'histogramme pour conserver toujours une bonne dynamique, même dans des zones de fort contraste. [4]

4.3 La suppression du bruit

C'est l'un des sujets les plus délicats du traitement des images. Il a vu couler beaucoup d'encre et de nombreuses méthodes lui ont été consacrées, tout d'abord très intuitives, mais progressivement de plus en plus complexes. [4]

5. Filtrage

On peut définir le filtrage comme le processus de remplacement d'un pixel par une valeur qui est fonction des données à proximité du pixel.

On distingue les filtres en deux grandes catégories (les filtres linéaires et les filtres non linéaires ou adaptatifs) :

5.1 Filtres linéaires

Le filtrage linéaire est un produit de convolution c-à-d une combinaison linéaire du voisinage du pixel concerné. Dans de nombreux cas on considère que l'image non dégradée $f(x, y)$ est affectée d'un bruit additif $n(x, y)$:

$$G(x, y) = f(x, y) + n(x, y)$$

Les différents types des filtres linéaires sont :

5.1.1 Filtre moyeneur (mean filtering ou box filtering)

Le principe est très simple : un pixel est remplacé par la moyenne de lui-même et de ses voisins. C'est dans la définition du voisinage que les filtres vont différer. On peut considérer un voisinage en connexité 4 ou 8, ou même encore plus large.

$$\begin{array}{cc} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} / 5 & \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} / 9 \\ \text{Connexité 4} & \text{Connexité 8} \end{array}$$

Pour une implémentation plus rapide, on préférera utiliser des filtres avec des coefficients entiers puis diviser ensuite le résultat par la somme des coefficients (cela permet d'effectuer des opérations sur des entiers plutôt que sur des réels).

Le filtre moyeneur est un filtre passe-bas permettant ainsi d'éliminer les hautes fréquences, correspondant au bruit. Son inconvénient est qu'il élimine également les hautes fréquences correspondantes aux détails de l'image : il rend ainsi l'image moins bruitée mais plus floue.

[1]



Image originale



Image filtrée par un moyeneur

Figure 1.5: Exemple du filtre moyen

5.1.2 Filtre gaussien (gaussian filtering)

Le principe de ce filtrage est une convolution avec une gaussienne.

Nous rappelons l'expression d'une gaussienne en deux dimensions, de moyenne nulle (x est un vecteur de dimension 2) :

$$G_0(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

Pour effectuer une convolution avec une gaussienne, on utilise un masque de convolution obtenu par discrétisation d'une gaussienne sur un noyau généralement de taille (2p+1)*(2p+1). certains masques sont à coefficients entiers pour permettre des calculs plus rapides, voire à coefficients puissances de deux (une multiplication ou division par 2 d'un entier revient à un décalage de 1 des bits le composant). [1]

La discrétisation de ce filtre pour un σ égale à 0.6 donne le masque suivant :

1/16 X

1	2	1
2	4	2
1	2	1

Masque du filtre

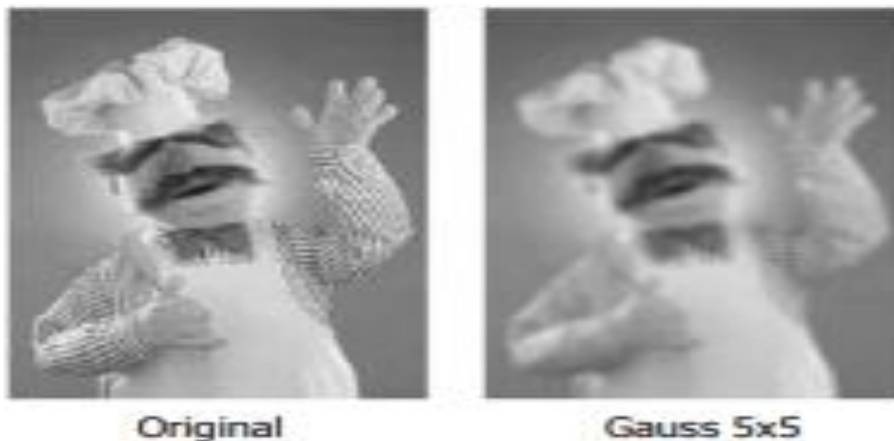


Figure 1.6 : Application du filtre gaussien

5.2 Filtres non linéaires

C'est un domaine qui est extrêmement vaste. Les opérateurs les plus courants en prétraitement d'image sont les filtres médians et les filtres morphologiques :

5.2.1 Filtre médian (median filter)

Le filtre médian est une amélioration du filtre conservatif. Son principe consiste à remplacer un pixel par la médiane de ses voisins. Ainsi, même si plusieurs pixels voisins sont bruités, on peut corriger le pixel courant. Ce filtre induit cependant un lissage puisque même des pixels corrects peuvent être modifiés. De plus, ce filtrage est plus coûteux car nécessite d'effectuer un tri des voisins pour chaque pixel. Plus le voisinage considéré est grand, plus l'algorithme

sera coûteux. On pensera donc, lors de l'implémentation, à utiliser un algorithme de tri rapide pour le nombre de pixel considéré.

Le filtre médian permet d'obtenir de bons résultats sur du bruit poivre et sel mais est aussi peu performant que le filtre conservatif pour le bruit gaussien. [1]

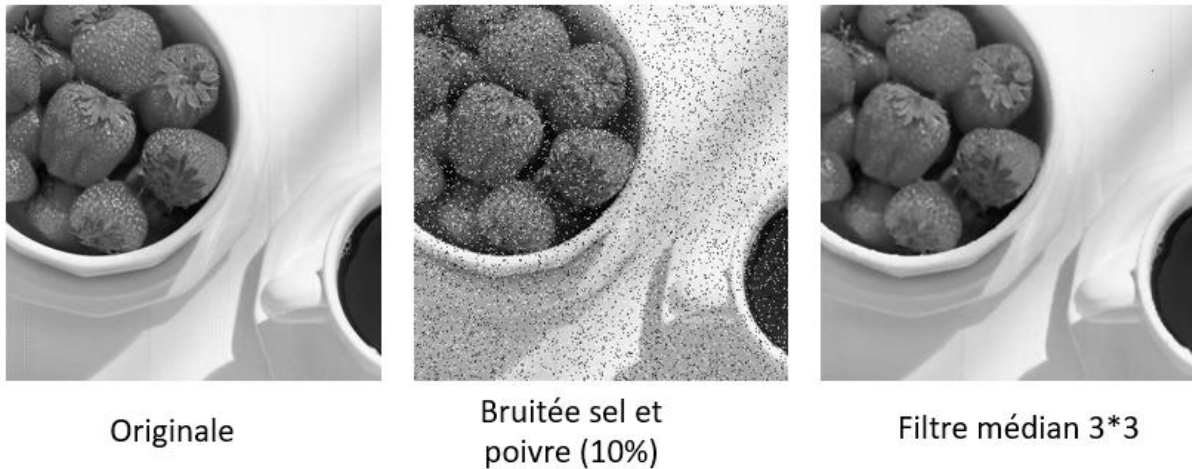


Figure1.7 : Principe du filtre médian

5.2.2 Filtres morphologiques

Les filtres morphologiques consistent à atténuer ou supprimer le bruit pour retrouver l'information utile sous-jacente. L'ouverture et la fermeture morphologiques filtrent le bruit quel que soit son amplitude relative, pourvu qu'elle soit positive pour la première et négatif pour la seconde.

Parmi les filtres morphologiques, le filtre alterné séquentiel est un de ceux qui ont les meilleures propriétés de réduction de bruit.

6 .Segmentation

6.1 La définition

La segmentation d'image est l'opération de partition d'une image à des composantes séparées ou à des objets séparés. Son but est de :

- Démarrer l'interprétation de l'image,
- Réduire la quantité de l'information dans l'image, sans nuire de façon sensible à la compréhension,
- Structurer les informations de l'image.

Dans la segmentation d'images, on distingue deux approches :

- Segmentation par régions,
- Segmentation par extraction de contours.

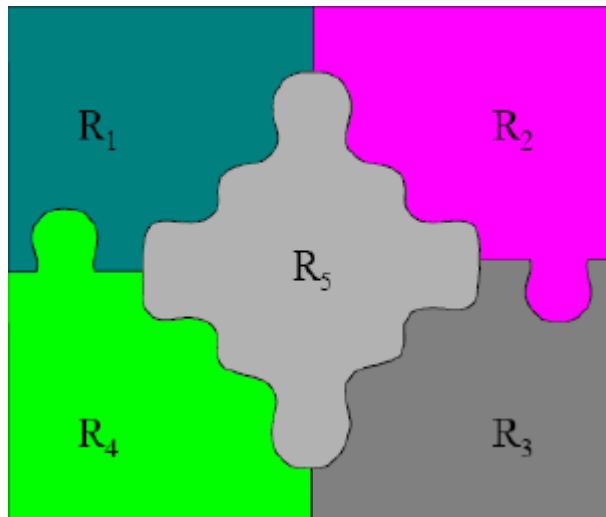


Figure 1.8 : Exemple de segmentation d'image

6.2 Définition formelle de la segmentation

Formellement, la segmentation d'une image A en régions R_i , $i= 1..n$, est définie par les propriétés suivantes :

1. $\cup_{i=1}^n R_i = I$
2. $R_i \cap R_j = \emptyset; \forall i, j$ telle que $i \neq j$
3. $p(R_i) = \text{vrai}; \forall i \in \{1, 2, \dots, n\}$
4. $p(R_i \cap R_j) = \text{faux}; \forall i, j$ telle que $i \neq j$
5. R_i est un composant connexe; $\forall i \in \{1, 2, \dots, n\}$

- P est un prédicat d'homogénéité.
- La première condition indique que l'union des régions nous ramène à l'image de départ.
- La deuxième indique que deux régions différentes sont disjointes.
- La troisième exprime que les pixels appartenant à une région doivent satisfaire le critère d'homogénéité.
- La quatrième exprime le critère d'homogénéité pour la segmentation en régions disjointes. [8]

6.3 Les méthodes de segmentation

La segmentation peut être abordée de deux points de vue dans la mesure où une région peut être définie par l'ensemble des pixels la composant (approche région de la segmentation) ou bien par les contours de la région (approche contour de la segmentation).

Ces deux approches sont duales du fait que chaque région possède un contour et qu'un contour délimite forcément une région.

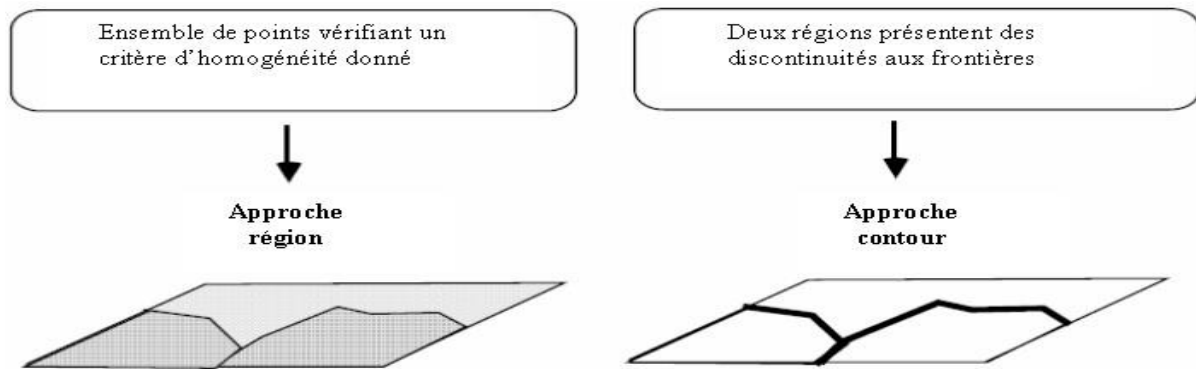


Figure 1.9 : Approche région et approche contour

On distingue deux méthodes principales :

6.3.1 Segmentation par régions

La segmentation d'image par l'approche région consiste à découper l'image en régions. Les pixels adjacents sont regroupés en régions distinctes selon un critère d'homogénéité ou de similarité donnée. Ce critère peut être, par exemple, le niveau de gris, couleur, texture...etc.

Un processus de groupement est répété jusqu'à ce que tous les pixels dans l'image soient inclus dans des régions. Cette approche vise, donc, à segmenter l'image en se basant sur des propriétés intrinsèques des régions.

Il existe plusieurs méthodes telles que la segmentation par croissance de région, par division de région, et par fusion de région que nous présentons ci-dessous.

6.3.1.1 croissance de région

Cette technique consiste à faire progressivement accroître les régions autour de leur point de départ. L'initialisation de cette méthode consiste à considérer chaque pixel comme une région. On va essayer de les regrouper entre elles avec un double critère de similarité des niveaux de gris et d'adjacence. Le critère de similarité peut par exemple être : la variance des niveaux de gris de la région R est inférieure à un seuil.

Le principe de l'agrégation de pixel est le suivant : on choisit un germe (Le point de départ est le choix d'un ensemble de pixels appelés « germes ») et on fait croître ce germe tant que des pixels de son voisinage vérifient le test d'homogénéité. Lorsqu'il n'y a plus de pixels candidats dans le voisinage, on choisit un nouveau germe et on itère le processus. [9]

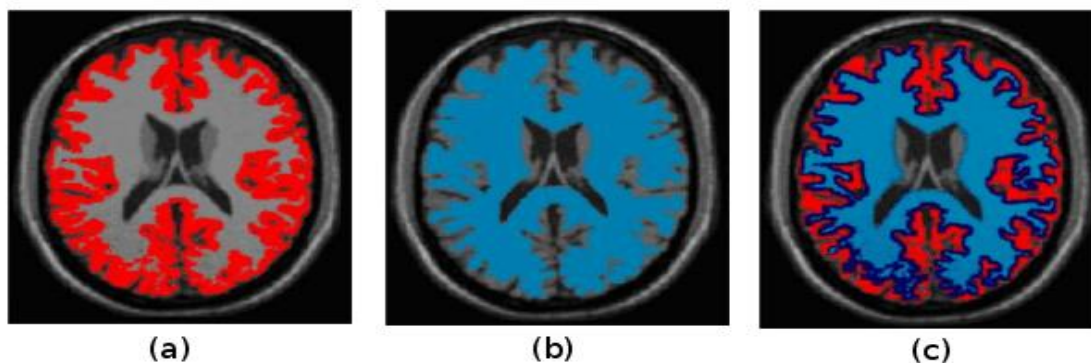


Figure 1.10 : Segmentation par croissance de régions

Parmi les avantages de cette technique, nous pouvons citer :

- La simplicité et la rapidité de la méthode.
- La segmentation d'objet à topologie complexe.
- La préservation de la forme de chaque région de l'image.

Cependant, il existe plusieurs inconvénients comme :

- L'influence du choix des germes initiaux et du critère d'homogénéité sur le résultat de la segmentation.
- Une mauvaise sélection des germes ou un choix du critère de similarité mal adapté peuvent entraîner des phénomènes de sous-segmentation ou de sur-segmentation.
- Il peut y avoir des pixels qui ne peuvent pas être classés.

6.3.1.2 Segmentation par fusion de régions

Les techniques de réunion (*region merging*) sont des méthodes ascendantes où tous les pixels sont visités. Pour chaque voisinage de pixel, un prédicat P est testé. S'il est vérifié les pixels correspondants sont regroupés dans une région.

Les inconvénients de cette méthode se situent à deux niveaux :

- Cette méthode dépend du critère de fusion qui peut influencer sur le résultat final de la segmentation.
- Elle peut introduire l'effet de sous-segmentation.

6.3.1.3 Segmentation par division de régions

La division consiste à partitionner l'image en régions homogènes selon un critère donné. Le principe de cette technique est de considérer l'image elle-même comme région initiale, qui par la suite est divisée en régions.

Le processus de division est réitéré sur chaque nouvelle région (issue de la division) jusqu'à l'obtention de classes homogènes. [8]

Cette méthode présente un inconvénient majeur qui est la sur-segmentation. Toutefois, ce problème peut être résolu en utilisant la méthode de division-fusion que nous présentons dans ce qui suit.

6.3.1.4 Segmentation par division –fusion

Ces méthodes combinent les deux méthodes décrites précédemment, la division de l'image en de petites régions homogènes, puis la fusion des régions connexes et similaires au sens d'un prédicat de regroupement. On part du principe que chaque pixel représente à lui seul une région. Deux régions seront fusionnées si elles répondent aux critères de similarité des niveaux de gris et d'adjacence de régions.

On s'arrête quand le critère de fusion n'est plus vérifié. [10]

Les inconvénients de cette méthode se situent à trois niveaux :

- Les régions obtenues ne correspondent pas, dans tous les cas, aux objets représentés dans l'image.
- Les limites des régions obtenues sont habituellement imprécises et ne coïncident pas exactement aux limites des objets de l'image.
- La difficulté d'identifier les critères pour agréger les pixels ou pour fusionner et diviser les régions.

6.3.2 Segmentation par contours

L'approche contour consiste à identifier les changements entre les régions. En général, un élément de contours est un point de l'image appartenant à la frontière de deux ou plusieurs objets ayant des niveaux de gris différents. Les variations d'intensité de lumière et de couleurs sont très bien perçues par le système visuel humain. la figure suivant montre quelques modèles de contours.

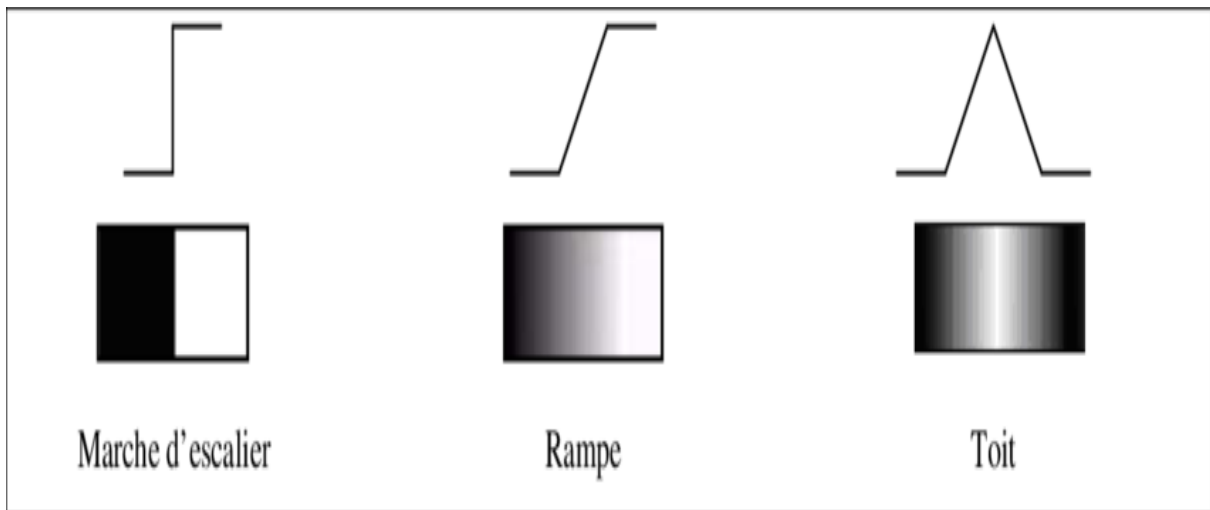


Figure 1.11 : les modèles de contours

- Marche d'escalier : le contour est net (contour idéal).
- Rampe : le contour est plus flou.
- Toit : il s'agit d'une ligne sur un fond uniforme.

Nous présentons dans ce qui suit les différentes méthodes adaptées pour la détection des contours dans des images en niveaux de gris. Pour ces dernières, deux familles de méthodes sont distinguées :

6.3.2.1 Méthodes dérivatives

Les méthodes dérivatives sont les plus utilisées pour détecter des transitions d'intensité par différenciation numérique (Première et deuxième dérivé). A chaque position, un opérateur est appliqué afin de détecter les transitions significatives au niveau de l'attribut de discontinuité choisi. Le résultat est une image binaire constituée de points de contours et de points non-contours.

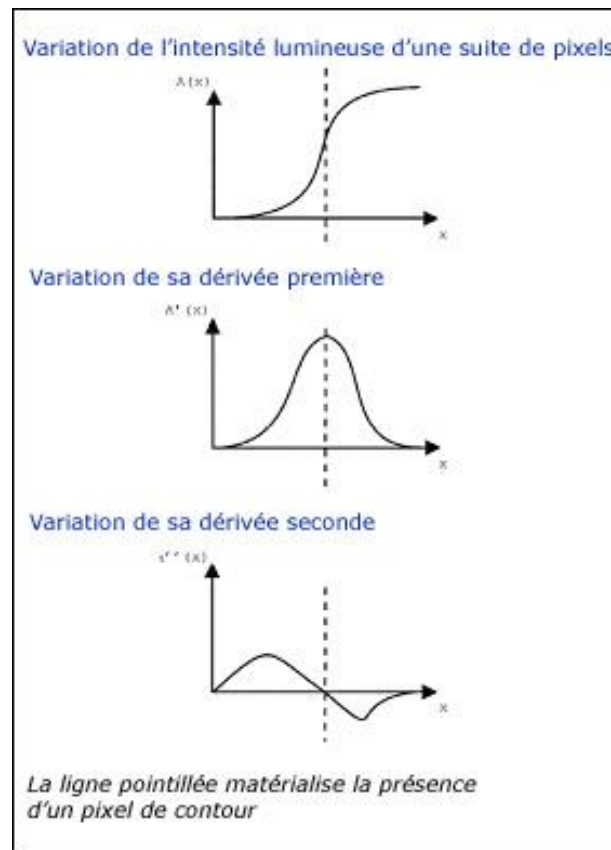


Figure 1.12 : contour et ses dérivées

De nombreuses techniques d'extraction de contours existent dans la littérature. Elles peuvent être classées comme suit [8]:

- Les algorithmes basés sur le gradient (ou opérateurs du premier ordre).
- Les algorithmes basés sur Laplacien (ou opérateurs du second ordre).

a) L'approche Gradient

On utilise la première dérivée pour calculer le gradient. Le gradient, en un pixel, est un vecteur caractérisé par une amplitude et une direction.

L'approximation des deux dérivées est donnée par la relation suivante :

$$\frac{\partial I(x,y)}{\partial x} = I(x + 1, y) - I(x, y), \frac{\partial I(x,y)}{\partial y} = I(x, y + 1) - I(x, y)$$

Le module :

$$G(x, y) = \|\nabla I(x, y)\| = \sqrt{G_x(x, y)^2 + G_y(x, y)^2}$$

La direction (argument) :

$$\arg G(x,y) = \arctan\left(\frac{G_y(x,y)}{G_x(x,y)}\right)$$

Il existe plusieurs opérateurs de gradient Parmi ses opérateurs, il y a les masques de Roberts, de Prewitt et de Sobel ...etc. [10]

➤ **Opérateurs de Sobel et Prewitt**

Les opérateurs de "Sobel" et de "Prewitt" permettent d'estimer la norme du gradient bidimensionnel d'une image en niveau de gris. Ces opérateurs consistent en une paire de masques de convolution 3x3.

Pour ces opérateurs les dérivées directionnelles horizontale et verticale s'expriment sous la forme : [11]

$$\frac{\nabla I}{\nabla J} = h_j * I(i, j) \text{ et } \frac{\nabla I}{\nabla I} = h_i * I(i, j)$$

Avec :

$$h_j = \begin{bmatrix} 1 & 0 & -1 \\ c & 0 & -c \\ 1 & 0 & -1 \end{bmatrix} \quad \text{et} \quad h_i = \begin{bmatrix} 1 & c & 1 \\ 0 & 0 & 0 \\ -1 & -c & -c \end{bmatrix}$$

Les matrices h1 et h2 sont appelées masques, Les masques de Prewitt sont définis par c=1 et les masques de Sobel par c=2.

➤ **Opérateur de Roberts**

Le détecteur de Roberts permet de calculer le gradient bidimensionnel d'une image de manière simple et rapide. Ce principe ne diffère pas beaucoup de celui des opérateurs de "Prewitt" et "Sobel". [10]

Les masques de convolution de Robert sont :

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \text{ et } \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

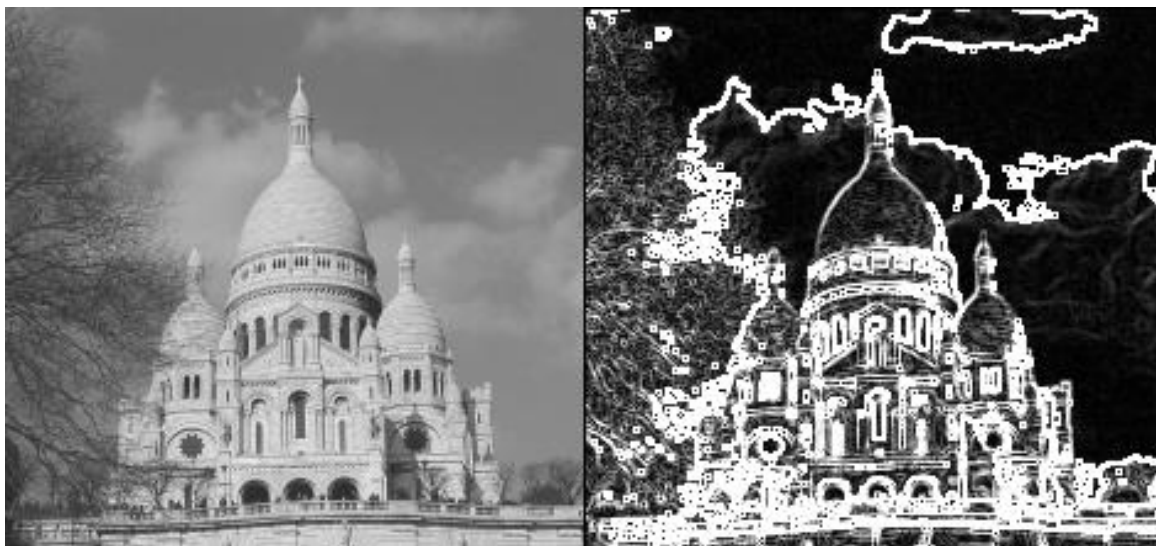


Figure 1.13 : Détection de contour

b) L'approche Laplacien

Dans cette approche, l'extraction des contours est basée sur le calcul des dérivées secondes, cela signifie que le Laplacien. [10]

$$\nabla f(i, j) = \frac{\partial^2 f}{\partial x^2}(i, j) + \frac{\partial^2 f}{\partial y^2}(i, j)$$

L'approximation des deux dérivées est donnée par la relation suivante :

$$\nabla^2 f = f(x + 1, y) + f(x - 1, y) + f(x, y - 1) + f(x, y + 1) - 4f(x, y)$$

La figure suivante représente la détection de contour en utilisant le Laplacien :



Figure 1.14 : contour détecté par Laplacien

6.3.2.2 Méthodes analytiques

➤ Approche de Canny et Deriche

Elle consiste à trouver un filtre optimal satisfaisant les 3 contraintes suivantes :

- **Une bonne détection** : faible probabilité d'oublier un vrai point de contour et une faible probabilité de marquer un point image comme contour alors qu'il ne l'est pas.
- **Une bonne localisation** : les points contours doivent être le plus près possibles de leur position réelle dans l'image.
- **Une réponse unique** : à un contour unique : un point de contour ne doit être détecté qu'une seule fois par le filtre mis en œuvre.

Le détecteur de contour de Canny est le plus utilisé. Il est basé sur trois critères : la détection (robustesse au bruit), la localisation (précision de la localisation du point contour), l'unicité (une seule réponse par contour). [12]

Au filtre de Canny, Deriche a proposé un autre filtre (condition initiale différente) qui permet une simplification de son implémentation nous préférons souvent le détecteur de Deriche, qui répond exactement aux mêmes critères de qualité que celui de Canny.[13]

6.3.2.3 Méthodes déformables

Les modèles déformables, introduits par Kass sont aussi connus sous les noms de « snakes » ou « contours actifs ».

L'intérêt principal des contours actifs est de détecter des objets dans une image en utilisant les techniques d'évolution de courbes. L'idée est de partir d'une courbe initiale, généralement un carré ou un cercle, et de la déformer jusqu'à obtenir le contour de l'objet.

En effet, celui-ci présente quelques inconvénients tels que la sensibilité à l'initialisation, au bruit, et le réglage difficile de ses différents paramètres. [14]

➤ Contours actifs

Les contours actifs tirent leur origine des modèles élastiques, mais la communauté s'accorde à les attribuer à l'équipe Kass, Witkin et Terzopoulos, qui introduisirent les Snakes ou courbes minimisantes. Les snakes tiennent leur nom de leur aptitude à se déformer comme des serpents. L'approche est variationnelle, c'est à- dire que l'équation d'évolution du contour actif se déduit de la minimisation d'une énergie modélisant l'objet d'intérêt. Depuis la publication de cette équipe, les modèles déformables sont devenus un sujet très important pour la communauté du traitement d'images. [9]

→ Principe de fonctionnement des contours actifs :

Un contour actif est une courbe définie dans l'image et qui est capable de se mouvoir sous l'influence de plusieurs forces qui vont la tirer ou la pousser vers la forme. Un snake consiste à placer aux alentours de l'objet à détecter une ligne initiale de contour, l'évolution se fait à travers un processus itératif qui déforme la courbe à chaque itération jusqu'à sa position finale

6.3.2.4 Les limites de segmentation par contour

Les principales limites des méthodes de détection de contour sont les suivantes : [8]

- Les contours extraits selon les méthodes classiques souvent ne correspondent pas nécessairement à la limite des objets. Dans de nombreuses images de basse qualité, quelques-unes des méthodes produisent des faux contours.
- Les techniques de détection de contour dépendent de l'information contenue dans le voisinage local de l'image. Il n'y a pas d'information globale.
- Après l'extraction des points de contours, ces derniers sont reliés afin de déterminer les frontières. Le processus de fermeture des contours peut parfois conduire à des discontinuités et des lacunes dans l'image.
- Il est souvent difficile d'identifier et de classer les contours parasites.

7. Conclusion

Dans ce chapitre, nous avons présentée une brève introduction aux concepts liés au domaine du traitement d'images. Les différentes définitions qui y sont développées sont celles des connaissances élémentaires de cette discipline, mais combien même elles sont essentielles pour L'initiation aux traitements approfondis des images.

Nous pouvons, à présent, aborder au prochain chapitre systèmes de recherche d'image par le contenu (CBIR).

CHAPITRE 02 :



**Systemes de recherche d'image par le
Contenu (CBIR)**

1. Introduction

L'expression « recherche d'images par le contenu » (« Content-Based Image Retrieval, CBIR, en Anglais) remonte aux travaux de Kato en 1992. Comme son nom l'indique, le principe de cette méthode est d'identifier des images à partir de leur contenu (c'est à dire à partir des données de l'image elles même et non à partir du texte associé aux images). Son système, ART MUSEUM, permet de retrouver des images d'art par couleurs et contours dont l'utilisateur a besoin. Le terme s'est étendu par la suite à tout procédé permettant de rechercher des images selon des descripteurs, qui peuvent être de type « signal », comme la couleur et la forme, mais peuvent également être des symboles.

Il existe plusieurs systèmes de recherche d'images (et donc les collections d'images). Elles incluent des applications judiciaires : les services de police possèdent de grandes collections d'indices visuels (visages, empreintes) exploitables par des systèmes de recherche d'images [15].

Les applications militaires sont certainement les plus développés, parmi ces applications on peut noter l'identification d'appareils ennemis sur les images radar, l'identification de cible à partir d'images satellites et les systèmes de guidage. Il en existe certainement d'autres mais le secret militaire ne permet pas d'en connaître les détails. Le journalisme et la publicité sont également d'excellentes applications. Les agences de journalisme ou de publicité maintiennent en effet de grosses bases d'images afin d'illustrer leurs articles ou supports publicitaires. Cette communauté rassemble le plus grand nombre d'utilisateurs de recherche par le contenu (davantage pour les vidéos) mais l'aide apportée par ces systèmes n'est absolument pas à la hauteur des espoirs initiaux [15].

Il existe également des demandes dans le domaine de l'architecture et de la construction : les architectes ont souvent besoin de se référer à des créations antérieures présentant des caractéristiques techniques similaires. Dans le domaine médical, en particulier en imagerie médicale, la recherche d'images par contenu peut aider à diagnostiquer en comparant une radiographie avec des cas précédents lui étant proche. Et bien d'autres applications dans le domaine de l'art, de l'éducation ou de loisirs.

Dans les difficultés suivantes sont identifiées : [16]

1. Comprendre les besoins des utilisateurs d'images et leurs comportements : de quoi les utilisateurs ont-ils besoin ?
2. Identifier une manière « convenable » de décrire le contenu d'une image. C'est une tâche rendue difficile par la subjectivité intrinsèque aux images.
3. Extraire des « descripteurs » des images brutes.
4. Pouvoir stocker de manière compacte un grand nombre d'images.
5. Comparer requêtes et images stockées de manière à refléter les jugements de similarité humains.
6. Accéder efficacement aux images par leur contenu.
7. Fournir des interfaces utilisables.

Donc on va d'abord démontrer les approches existantes ainsi que leurs limitations. Nous commençons par citer les composants d'un système de recherche d'images par le contenu, puis la taxonomie des systèmes selon leur niveau d'abstraction.

2. Composants d'un CBIR

Ici on décrit les caractéristiques communes à la plupart des approches : le traitement de la base d'images, les requêtes mise en correspondance puis la présentation des résultats et la (figure 2.1) résume ces étapes : [15]

Dans un premier temps (2), des descripteurs sont calculés à partir de chaque image de la collection (1), ils peuvent être de type signal ou/et symbolique (le vocabulaire d'indexation). Les données extraites (à présent représentatives du contenu de l'image du point de vue du système) constituent la base d'index (3). Les requêtes de l'utilisateur (4) sont alors transformées afin d'être comparables avec la base d'index (5) ; une mise en correspondance (6) entre la requête transformée et la base d'index permet ensuite de produire le résultat de la requête (7).

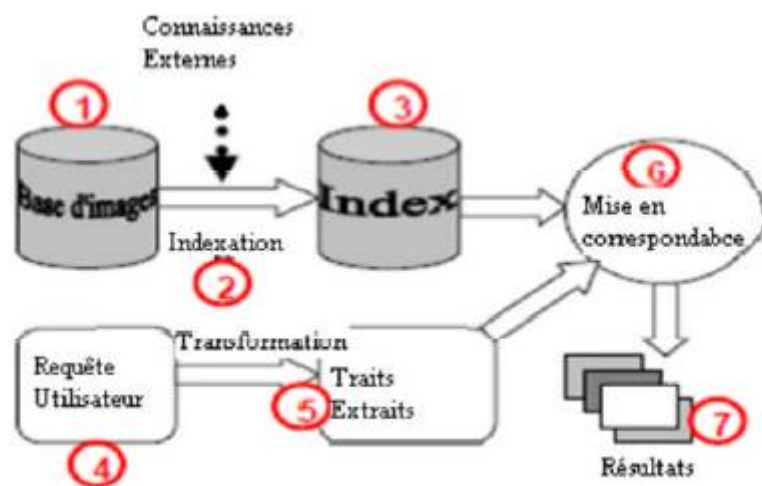


Figure 2.1 : Principaux composants d'un Système de Recherche d'images par le Contenu

2.1 Les bases d'images utilisées

La collection (ou base) d'images est la donnée principale du CBIR c'est une base de données qui regroupe de plusieurs centaines de milieu d'images. Ces bases d'images possèdent déjà des classes définies où chaque image n'appartient qu'à une seule classe.

2.1.1 La base de Wang

La base d'images de Wang est un sous-ensemble de la base d'images Corel. Cette base d'images a été créée par le groupe du professeur Wang de l'université Pennsylvania State et est disponible à l'adresse : <http://wang.ist.psu.edu/>. La base originale contient 1000 images naturelles en couleurs, divisées en 10 classes, chaque classe contient 100 images. L'avantage de cette base est de pouvoir évaluer les résultats.

(La figure 2. 2) présente un exemple de chaque classe :



Figure 2. 4 : Les objets utilisés dans COIL-20 (Deselaers, 2003)

2.1.3 Pollen

La base de 347 images de grains de pollen (31 classes au total) de l'INRIA Sophia-Antipolis et provenant du projet Européen ASTHMA1 (Bonton, 2002) [17]. Quelques exemples de cette base d'images sont montrés dans (la figure 2.5).

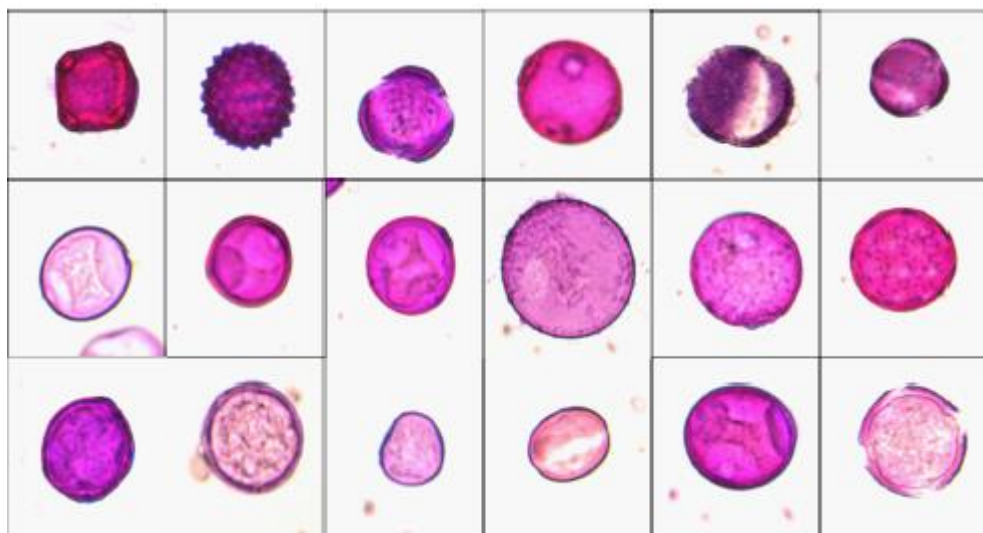


Figure 2.5 : Quelques images exemples de la base de Pollens

2.1.4 CUREt (Columbia Utrecht Reflectance and Texture Database)

Cette base a plus de 14000 images de textures (70 textures x ~200 images/texture) des Universités Columbia et d'Utrecht (CUREt). [17]

Quelques exemples de cette base d'images sont montrés dans (la figure 2.6).

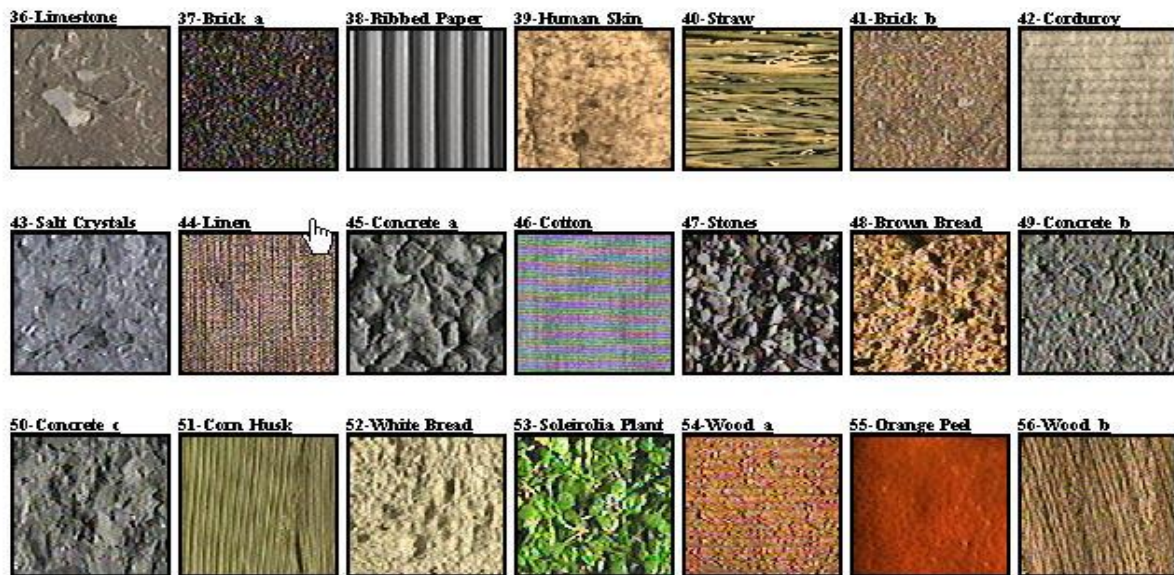


Figure 2.6 : Quelques images exemples dans la base de CURET

2.1.5 La base de FeiFei

Cette base contient des images de 101 objets collectées par Fei-Fei Li, Marco Adreetto et Marc Aurolio Ranzato. Chaque objet à de 40 à 800 images ont été prises, de taille de 300×200 pixels pour chaque images. [17]

(La figure 2.7) montre quelques exemples de cette base :



Figure 2.7 : Quelques images exemples dans la base de Fei-Fei

2.2 L'indexation

Depuis quelques années, l'indexation d'images a émergé comme nouveau thème de recherche dans le domaine de l'analyse des images numériques. C'est l'étape la plus délicate de la recherche d'image à partir d'une collection d'images (qui peut être spécifique, hétérogène..).C'est un ensemble de processus utilisés pour extraire des informations à partir de documents image et représenter ces informations dans une structure appelée index, pour la caractériser le mieux possible et de ne travailler que sur ce modèle lors de la recherche. Cela permettra une meilleure organisation des données, de limiter la quantité de données examinées durant une recherche, d'y accéder rapidement et de confiner la recherche au maximum.

L'indexation peut être **générique**, pouvant caractériser des collections hétérogènes, ou **spécifique**, adaptée à un type d'image particulier.

Un système d'indexation comprend généralement deux phases de traitement à savoir, l'indexation logique et l'indexation physique :

- **indexation logique**

qui consiste à extraire et à modéliser les caractéristiques de l'image (forme, couleur, texture).chacune de ces caractéristiques pouvant être considérée pour une image entière ou pour une région de l'image.

- **indexation physique**

Elle détermine une structure efficace d'accès au données afin de trouver rapidement une information. De nombreuses techniques basées sur des graphes sous formes d'arbres ont été proposées. Enfin l'indexation varie d'un système à l'autre par son niveau d'abstraction : l'extraction de l'histogramme couleur est une opération simple, tandis que l'identification des personnes ou des objets est beaucoup plus compliquée et nécessite un apprentissage préalable.

2.3 La gestion des index

C'est la manière de gestion d'index des images : stockage et accès. La gestion des index, anecdotique d'une collection de taille moyenne, devient un problème indispensable lorsqu'on travaille sur une grande base de taille conséquente. La méthode la plus basique de stockage des index est la liste séquentielle, en mémoire ou dans un fichier. Cependant lorsque le nombre d'images augmente, le temps d'accès au images augmente linéairement, il est généralement nécessaire d'organiser les index de manière hiérarchique, en la forme d'un arbre (organiser selon des descripteurs ou , par exemple, un tableau de « hash codes »,afin d'accélérer la vitesse d'obtention des information .

2.4 Les requêtes

La constitution de la requête est la première étape de la recherche d'images. Elles permettent au système de retrouver les images désirées par l'utilisateur. Plusieurs requêtes sont proposées suivant les besoins de l'utilisateur et le type de base de données images, requête par esquisse (croquis), Requête par mots clés, requête par exemple et la combinaison de celles-ci afin d'accéder à un niveau d'abstraction supérieur.

2.4.1 Requête par mots clés

La plupart des systèmes de recherche d'images développés utilisent des mots clés ou des descripteurs textuelles pour caractériser chaque image de la base (ex : recherche d'images sur Internet). Ces images sont recherchées suivant un ou plusieurs critères, par exemple trouver les images contenant 80% de rouge. Donc, le système se base sur l'annotation manuelle et textuelle d'images. Beaucoup de moteurs de recherche d'images tels que Google, Yahoo...utilisent cette façon. Cette méthode n'est pas parfaite parce que quelques mots qui n'expriment pas le sens d'une image.

(La figure 2.8) donne un exemple, l'utilisateur veut trouver des images contiennent une (des) voiture(s) avec le ciel cependant les premières images résultats ne sont pas pertinentes car ils n'ont ni les voitures ni le ciel. De plus l'indexation de ces images représente une tâche longue et répétitive pour l'humain, surtout avec les bases d'images qui deviennent aujourd'hui de plus en plus grandes. Elle est subjective à la culture, à la connaissance et aux sentiments de chaque personne. [17]

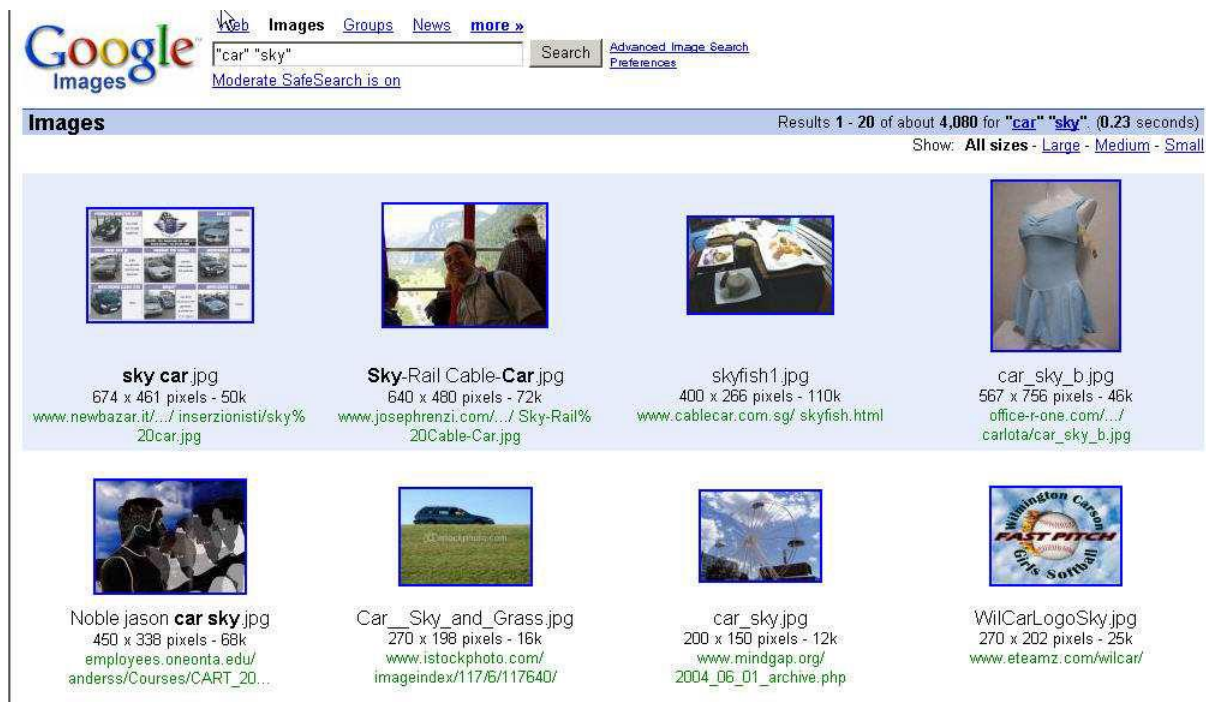


Figure 2.8 : Un exemple de recherche d'images dans Google

Puisque la requête par mots clés a des désavantages, les systèmes de recherche d'images par le contenu utilisent souvent les trois autres façons :

Requête par exemple, requête par esquisse et requête par contenu.

2.4.2 Requête par esquisse

Le système fournit à l'utilisateur des outils lui permettant de constituer une esquisse (dessin) correspondant à ses besoins. L'esquisse fournie sera utilisée comme exemple pour la recherche. L'esquisse peut être une ébauche de forme ou contour d'une image entière ou une ébauche des couleurs ou textures des régions d'une image. L'utilisateur choisira, en fonction de la base d'images utilisée, de ses besoins et préférences, l'une ou l'autre de ces représentations. Cette technique présente l'inconvénient majeur qu'il est parfois difficile pour l'utilisateur de fournir une esquisse malgré les outils qui lui sont fournis. [18]

2.4.3 Requête par le contenu

Cette technique se base sur l'identification des images à partir de leur contenu, et pas du texte associé aux images. L'indexation des images, et leurs paramètres qui sont la couleur, la texture, l'intensité ou, encore, les formes contenues dans l'image doivent être extraits au préalable. et ils fournissent une "signature".

2.4.4 Requête par l'exemple

Ici, l'image peut être soit fournie par l'utilisateur, soit être choisie par ce dernier dans la base d'images utilisée. L'utilisateur utilise une image (ou une partie d'image) qu'il considère similaire aux images qu'il recherche. Cette image est appelée image exemple ou requête. Cette méthode ne nécessite pas de connaissances approfondies pour manipuler le système.

2.5 Analyse de la requête

Le but de cette étape est de convertir la demande de l'utilisateur afin qu'elle soit comparable au index de la base d'images ; par conséquent, il s'agit généralement d'extraire le même type de descripteurs que les descripteurs extraits de la base d'images lors de l'indexation.

2.6 Mise en correspondance requête / base

Il s'agit d'estimer dans quelle mesure une image (son index) satisfait une requête donnée. Dans le contexte de la recherche d'images, cela se ramène souvent à calculer la similarité entre les caractéristiques extraites de la requête et les caractéristiques de chaque image dans la base. Cela aboutit généralement à une valeur de correspondance qui caractérise la pertinence (du point de vue du système) d'une image par rapport à la requête. Cette mise en correspondance peut être simple ou complexe [15]. La phase de mise en correspondance peut également inclure une pondération des descripteurs. Ponderer les descripteurs permet d'éliminer une partie du bruit dans la mesure où les descripteurs les moins pertinents voient leur influence diminuer dans l'évaluation de la similarité requête/image [15].

La mise en correspondance peut également inclure un bouclage de pertinence. Le but est également d'éliminer le bruit (augmenter la précision) en tentant de converger vers une précision maximale [15].

2.7 La présentation des résultats

Les résultats de la requête sont affichés sous forme de listes d'images (réduite en vignettes), et ces listes sont classées par ordre décroissant de pertinence. Par rapport aux documents textuels, l'avantage des images est que vous pouvez voir l'intégralité du document en un coup d'œil, ce qui vous permet de visualiser un grand nombre de résultats et d'effectuer des comparaisons plus rapidement. Comme mentionné ci-dessus, la présentation des résultats est généralement associée à la possibilité d'interaction, ce qui permet, par exemple, d'optimiser la requête en indiquant des résultats pertinents et non pertinents (cycles pertinents) au système, permettant ainsi une reformulation automatique de la requête.

3. Mesures pour évaluer un système

Avant d'exécuter un système de recherche d'information, une évaluation qui permet de mesurer les performances de ces systèmes est nécessaire. Le temps de réponse et l'espace utilisé sont les mesures les plus courantes pour évaluer un système. Plus le temps de réponse est court, plus l'espace utilisé est petit, et plus le système est considéré bon. Mais en plus de ces deux méthodes de mesure, on s'intéresse à d'autres mesures. Dans le système de recherche d'informations, l'utilisateur s'intéresse aux réponses pertinentes du système. Donc les systèmes de recherche d'informations exigent l'évaluation de la précision de la réponse.

Le système d'indexation et de recherche d'images est un système de recherche d'informations. Dans les systèmes de recherche d'images, les auteurs ont souvent utilisé les mesures d'évaluation pour évaluer des systèmes de recherche d'informations [19].

Dans cette partie, nous allons présenter les deux paramètres les plus courantes: le rappel et la précision. Ces mesures sont interdépendantes. Nous utilisons donc souvent une courbe de rappel et précision. Ensuite, nous introduisons d'autres mesures que nous utilisons également pour évaluer un système de recherche d'informations.

3.1 Rappel et précision (en anglais : Recall and Precision)

- **Le rappel**

Le rappel est défini par le nombre d'image pertinents retrouvés au regard du nombre d'image pertinents que possède la base d'image.

Le rappel est donc calculé comme suit :

$$\text{Rappel} = \frac{|Ra|}{|R|}$$

- **La précision**

La précision est le nombre d'image pertinents retrouvés rapporté au nombre d'image total proposé pour une donnée.

$$\text{Précision} = \frac{|Ra|}{|A|}$$

Où :

- _ I : une image requête
- _ R : l'ensemble d'images pertinentes dans la base d'images utilisée pour évaluer.
- _ |R| : le nombre d'images pertinentes dans la base d'images.
- _ A : l'ensemble des réponses.
- _ |A| : le nombre d'images dans l'ensemble des réponses.
- _ |Ra| : le nombre d'images pertinentes dans l'ensemble des réponses.

Des définitions sont montrées dans (la figure 2.9):

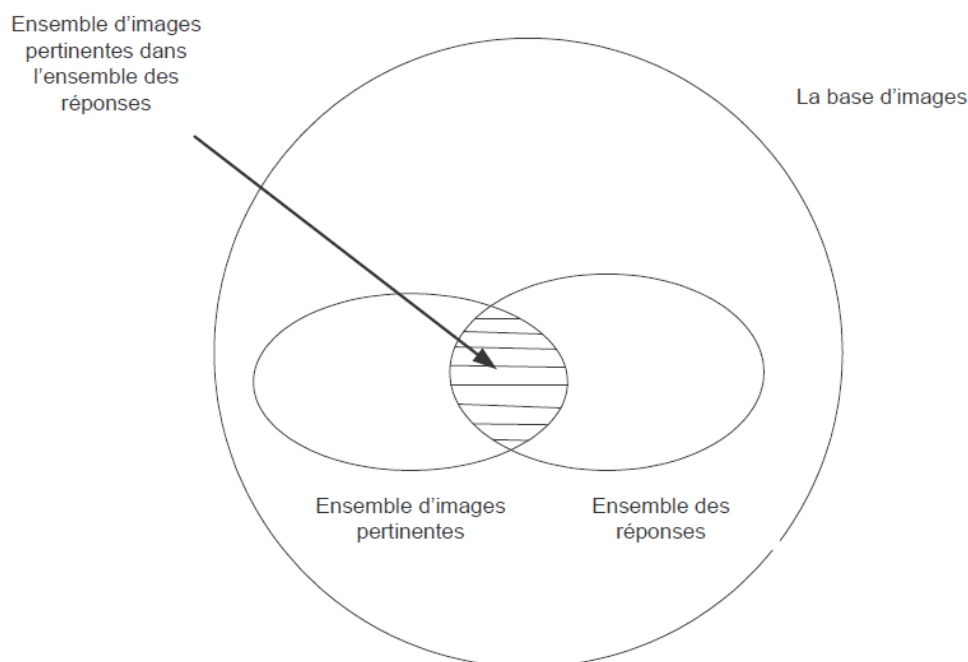


Figure 2.9 : Le rappel et la précision pour une requête (Yates, 1999)

L'exemple présenté dans (la figure 2.10) illustre le calcul de rappel et précision.



Calcul de Rappel et précision :

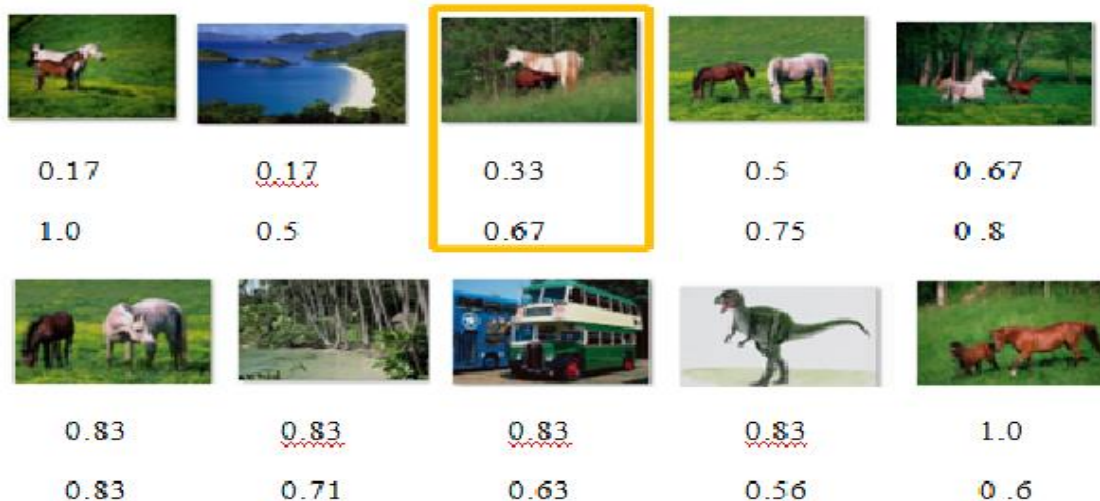


Figure 2.10 : Exemple de calcul rappel et précision

- Rappel=2/6=0.33
- Précision =2/3=0.67

3.2 La courbe de rappel et précision

Le rappel et la précision sont les mesures importantes, elles sont utilisées pour développer la performance de notre système. On donne souvent une distribution de rappel et précision sous forme d'une courbe, parce qu'une paire de ces deux valeurs ne peut pas indiquer la performance du système. C'est un calcul statistique sur la base de données mettant aussi en évidence la capacité de la méthode étudiée à retrouver les classes d'images.

➤ **D'autres mesures**

- **La moyenne harmonique**

La moyenne harmonique de rappel et précision $F(j)$ a été définie par la formule suivante :

$$F(j) = \frac{2}{\frac{1}{r(j)} + \frac{1}{p(j)}}$$

Où :

- _ $r(j)$: le rappel quand on donne j images en réponse.
- _ $P(j)$: la précision quand on donne j images en réponse.

La valeur de cette mesure est entre 0 et 1.

- Si la valeur est 0 ça veut dire qu'aucune image pertinente a été retrouvée.

-Si la valeur est 1, toutes les images pertinentes ont été retrouvées.

- **La mesure E**

La mesure E de rappel et précision $E(j)$ a été définie par la formule suivante :

$$E(j) = \frac{1+b^2}{\frac{b^2}{r(j)} + \frac{1}{p(j)}}$$

Où :

- _ $r(j)$: le rappel quand on donne j images en réponse.
- _ $P(j)$: la précision quand on donne j images en réponse.
- _ b est le paramètre qu'on peut choisir.

4. Conclusion

Dans ce chapitre nous avons décrit les concepts de base pour la construction d'un système de recherche d'images par le contenu. Nous avons abordé les différents composants d'un CBIR, qui sont les bases d'images utilisées, les différents types de requêtes et l'indexation. Puis nous avons parlé de la représentation des images dans un CBIR.

A la fin de ce chapitre nous avons présenté les différentes mesures pour évaluer un système de recherche d'image. la recherche d'images par le Contenu s'est imposée dans tous les domaines de notre vie quotidienne, elle attire beaucoup d'attention et devient un axe de recherche très actif avec tous ses aspects.

Dans le chapitre qui suit, nous allons présentés les différents descripteurs extraient a partir d'une image et les mesures de similarité entre ces descripteurs.

CHAPITRE 03 :



Descripteurs d'image et mesures de similarité

1. Introduction

Aujourd'hui, dans différents domaines, nous utilisons de plus en plus le contenu visuel comme support de communication. Grâce à la densité et la richesse de leur contenu, l'image et la vidéo numérique font partie intégrante des systèmes d'indexation et recherche d'image par le contenu. Ces systèmes permettent de rechercher les images d'une base d'images en fonction de leurs caractéristiques visuelles. La représentation du contenu des données sources mises sous forme de vecteur caractéristique. L'information obtenue dans cette étape est une sorte de résumé des images de la base (segmentation en régions, couleur, texture, relations spatiales,...). La transformation est généralement gourmande en temps de calcul.

L'objectif de ce chapitre est de présenter les principales caractéristiques d'images qui sont encore appelées caractéristiques de bas-niveau, qui sont la couleur, la texture, la forme et tout autre caractéristique de l'image qu'on peut imaginer. Et puisque La recherche d'images se fait par comparaison de ces caractéristiques nous présenterons les mesures de similarité entres les différents descripteurs d'images ,qu'elles se fait à l'aide des différents distances comme la distance de Minkowski, distance quadratique et distance de Bhattacharya ,etc.

2. Descripteurs d'images

Pour décrire l'image, des descripteurs de bas-niveau sont utilisés, appelés aussi vecteurs caractéristiques, tels que la couleur, la texture et la forme...

2.1 Descripteurs de couleur

La couleur est l'information visuelle la plus utilisée dans les systèmes de recherche par le contenu. Elle forme une partie significative de la vision humaine. La couleur est devenue la première signature employée pour la recherche d'images par le contenu en raison de son invariance par rapport à l'échelle, la translation et la rotation [15]. Un descripteur couleur repose sur deux choix : la sélection d'un espace de couleur et le choix d'une représentation.

2.1.1 Les espaces de couleurs

Une couleur est généralement représentée par trois composantes. Ces composantes définissent un espace de couleurs. On parle de l'information couleur, ou intensité lumineuse et chromaticité. Il existe plusieurs espaces colorimétriques qui ont chacun certaines caractéristiques intéressantes, on distingue les espaces de couleurs définis selon des propriétés comme RGB (Red, Green, Blue), et ceux basés sur la perception humaine des couleurs comme HSV (Hue, Saturation, Value).

2.1.1.1 Espace RVB ou RGB en anglais

C'est l'espace de base le plus utilisé, car c'est celui employé par plusieurs appareils de capture d'images qui effectuent leurs échanges d'informations uniquement en utilisant les triplets (Rouge, Vert, Bleu). L'espace RVB est utilisé couramment dans tous les systèmes de vision automatiques, mais il n'est pas forcément le mieux adapté, car si on aura un simple

changement d'éclairage dans la scène modifié les 3 composantes et ses couleur ne sont pas toujours les mêmes [20].

(La figure 3.1) suivante montre une image originale et les trois images couleur dans l'espace RGB :



Figure 3.1 : Images couleurs dans l'espace RGB

2.1.1.2 Espace TSV

Le modèle TSV pour Teinte Saturation Valeur (HSV en anglais pour Hue saturation Value ou HSB pour Hue Saturation Brightness), est basé sur les trois composantes définies par une approche psychologique et perceptuelle de la couleur la teinte, la saturation et la luminance ou valeur. Le modèle TSV a été créé en 1978 par Alvy Ray Smith [21]. C'est une transformation non linéaire de l'espace de couleur RVB.

Le modèle TSV est particulièrement utilisé dans les applications graphiques. Dans beaucoup de contextes applicatifs, un utilisateur est amené à choisir une couleur qui sera appliqué à un élément. Utilisée de cette manière, la roue TSV est souvent employée [21].

La représentation conique est bien adaptée pour représenter tout l'espace TSV en un seul objet (Figure 3.2).

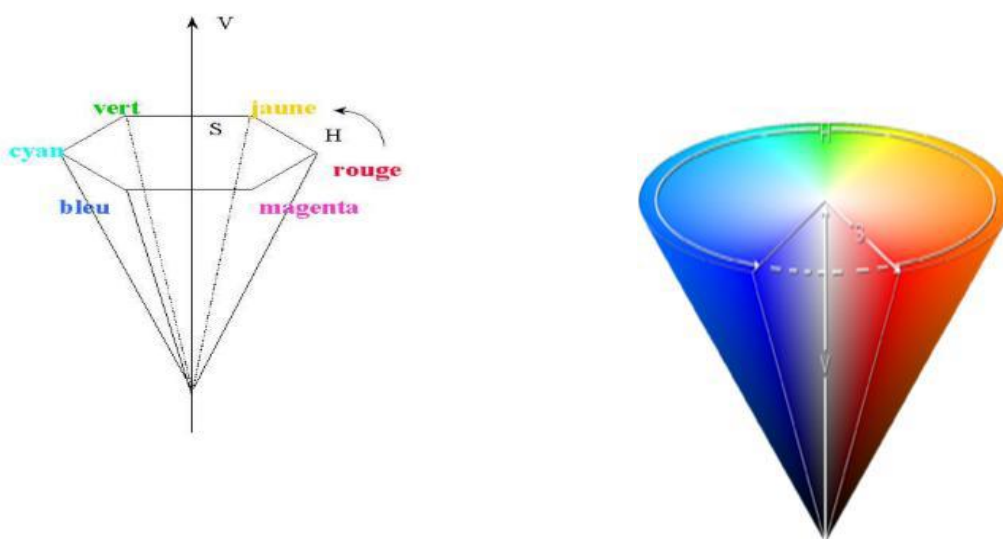


Figure 3.2 : Représentation conique de tout l'espace TSV en un seul objet.

2.1.1.3 Système XYZ

La CIE (Commission Internationale de l'Éclairage) est une organisation internationale chargée d'établir des normes et des recommandations reconnues par tous les pays, afin de pouvoir quantifier la couleur : c'est la base de la colorimétrie, science de la mesure de la couleur [15].

Le système XYZ a été établi par la CIE afin de pallier à certains inconvénients du système RGB.

L'espace CIE XYZ prend en compte la sensibilité de l'œil. La composante Y représente la luminance, X et Z contiennent l'information de chrominance. Il est rarement utilisé en recherche d'images car il n'est pas uniforme de point de vue humaine. De plus, il n'est pas facile d'interpréter les valeurs de tristimulus X, Y et Z et d'interpréter les couleurs qu'il représente [22].

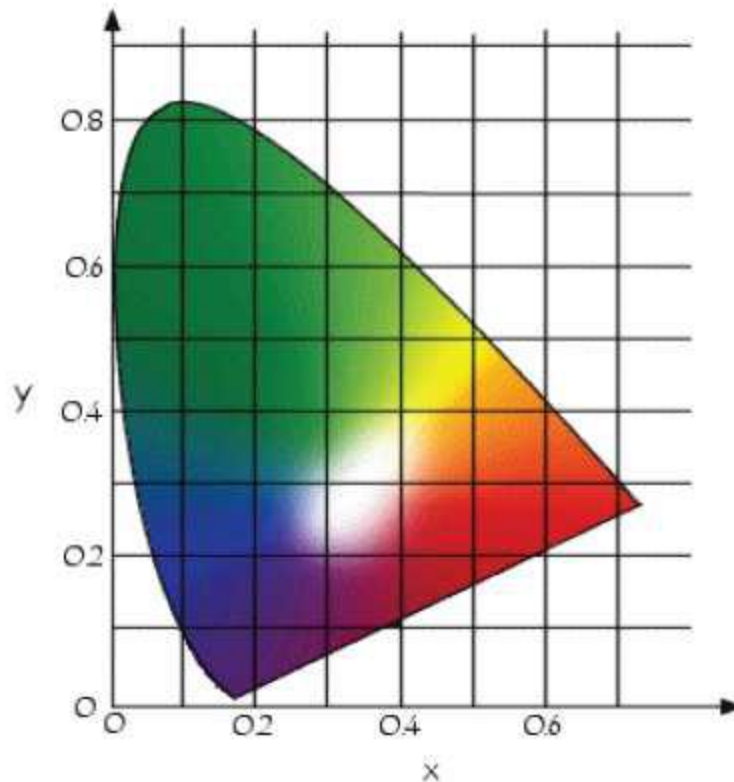


Figure 3.3 : Représentation de l'espace couleur XYZ

2.1.1.4 Système L*u*v

Les systèmes RGB et XYZ ne sont pas perceptuellement uniformes: une même distance euclidienne entre deux couleurs ne sera pas perçue de la même façon par l'être humain selon les couleurs [15]. Le système CIE-LUV, appelé aussi L*u*v, est un système perceptuellement uniforme que l'on peut obtenir à partir du système XYZ. C'est un espace de couleur pour la caractérisation des écrans. Il abandonne la linéarité pour représenter plus fidèlement les

écarts entre les couleurs perçues par la vision humaine et permet la mise en place d'une méthode de mesure des écarts de couleurs.

2.1.1.5 Système L*a*b

Le système lab est issu du CIE XYZ. il tient en compte la réponse logarithmique de l'œil. Il possède le grand avantage d'être uniforme. Cet espace est grandement utilisé dans les systèmes de comparaison d'images. Il est très utile dans le cas de mélanges de pigments, par exemple, pour l'industrie graphique ou du textile.

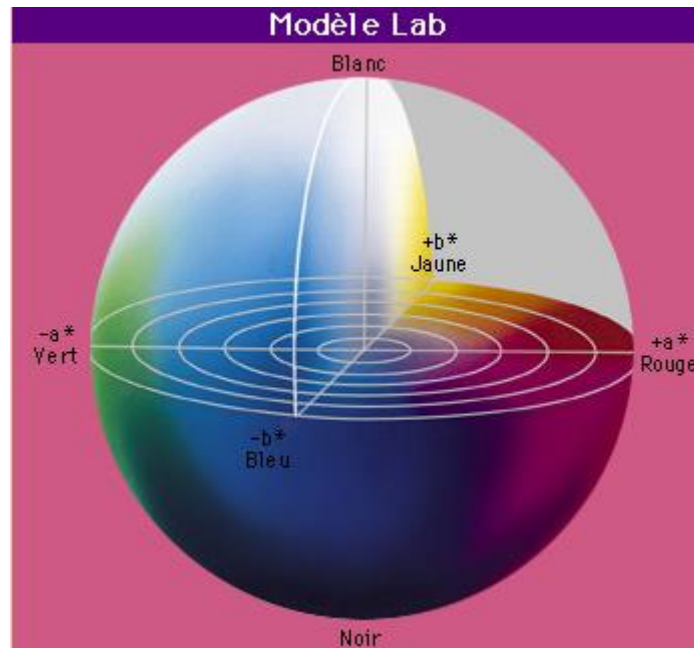


Figure 3.4 : Espace de couleur Lab

2.1.2 L'Histogramme

Les histogrammes sont des outils fondamentaux du traitement d'images avec de très nombreuses applications. Ils sont faciles et rapides à calculer, et robustes à la rotation et à la translation. Un histogramme est un outil statistique qui permet d'estimer la densité de probabilité d'une distribution à partir d'échantillons. Il est définie comme une fonction discrète qui associé à chaque valeur d'intensité le nombre de pixels prenant cette valeur , c'est-à-dire comptant le nombre de pixel pour chaque intensité (couleur) de l'image et pour mieux visualiser le descripteur de ces intensités de l'image.

(Figure 3.5).

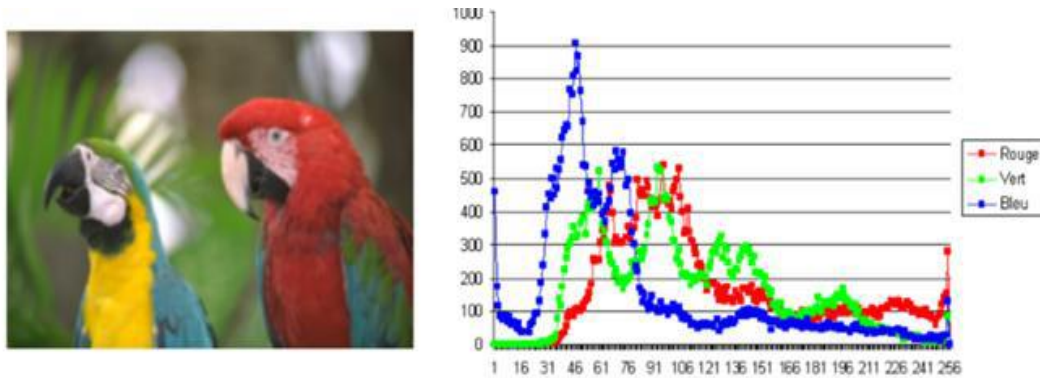


Figure 3.5 : Image couleur et son histogramme des trois composantes RVB

Cependant l'utilisation d'histogrammes pour l'indexation et la recherche d'images pose quatre problèmes : premièrement, ils sont de grandes tailles, donc il est difficile de créer une indexation rapide et efficace. Deuxièmement, ils ne possèdent pas d'informations spatiales sur les positions des couleurs. Troisièmement, ils sont sensibles à de petits changements de luminosité, ce qui pose un problème pour comparer des images similaires, et quatrièmement, ils sont inutilisables pour la comparaison partielle des images.

2.1.3 Les moments de couleurs

La méthode d'histogramme utilise la distribution complète de la couleur. On doit stocker de nombreuses données. Au lieu de calculer la distribution complète, dans les systèmes de recherche d'images, on calcule seulement des caractéristiques dominantes de couleur tel que l'espérance, la variance et d'autres moments. Dans [23] les auteurs ont prouvé que les méthodes des moments de couleurs utilisées marchent plus vite et donnent des résultats meilleurs que les méthodes d'histogrammes.

Les moments de couleur est une représentation compacte comparée aux autres descripteurs de couleur. Car seulement 9 valeurs (3 pour chaque composante chromatique) sont utilisées pour représenter le contenu d'une image. Pour cette raison ils peuvent diminuer le pouvoir de discrimination (description). [24]

2.1.4 Corrélogramme de couleurs

Les histogrammes donnent une bonne estimation de la densité de couleur d'une image, mais ils souffrent d'un manque d'information locale parce que ces derniers n'enregistrent pas la structure locale des couleurs mais seulement la composition globale, donc on peut perdre toute l'information sur la distribution spatiale des couleurs.

Donc les corrélogrammes ont été proposés pour l'indexation et la recherche d'images par le contenu, et semblent donner de très bons résultats en termes de reconnaissance ; ils enregistrent la probabilité de présence d'une couleur dans un voisinage prédéfini d'une autre couleur et qualifient non seulement la distribution de couleurs des pixels, mais aussi la corrélation spatiale entre les paires de couleurs. Ces informations sont représentées sous la forme d'un histogramme à trois dimensions : les deux premières dimensions représentent les combinaisons possibles de paires de pixels et la troisième dimension représente leurs distances spatiales [18].

Le corrélogramme est un tableau indexé par des paires de couleurs, où le *kième* entré (i,j) spécifie la probabilité de trouver un pixel de couleur j à une distance k d'un pixel de couleur i dans une image. Soit I l'ensemble de pixels d'une image et $I_c(i)$ l'ensemble de pixels dont la couleur est $c(i)$. [18]

Donc le corrélogramme de couleurs est défini par :

$$Y_{i,j}^{(k)} = \text{pr}_{p1 \in I_{c(i)}, p2 \in I} [p2 \in I_{c(i)}, |p1 - p2| = k]$$

Où $i,j \in \{1,2,\dots,N\}$, $K \in \{1,2,\dots,d\}$ et $|p1-p2|$, est la distance entre les pixels $p1$ et $p2$.

Cependant, l'implémentation des corrélogrammes est très délicate. Même si l'espace mémoire n'est pas forcément supérieur à l'espace utilisé par les histogrammes, leur création est très couteuse en temps de calcul, ce qui nécessite une implémentation très optimisée pour qu'ils soient utilisables avec un temps de calcul raisonnable.

2.1.5 Cohérence spatial

La cohérence spatiale a pour but de combler en partie, l'absence d'information spatio-colorimétrique de l'image dans le descripteur précédent. Ce descripteur est calculée pour chaque classe de couleur identifiée. Un histogramme de connexité est calculé par la formule suivante :

$$H_I(c) = \sum_{i=0}^{X=1} \sum_{j=0}^{X=1} \delta(I(i,j), c) \alpha(i,j)$$

I est l'image segmentée de taille (X,Y) , c 'est la couleur du pixel (i,j) , δ est le symbole de Kronecker et $\alpha(i,j)$ définie par :

$$\alpha(i,j) = \begin{cases} 1 & \text{si } \forall k, k' \in (-W, W) I(i+k, j+k') + I(i,j) \\ 0 & \text{sinon} \end{cases}$$

La fenêtre $(2W+1)*(2W+1)$ représentant le degré de compacité souhaité.

La cohérence spatiale est donnée par le rapport suivant :

$$SCR(c) = \frac{H_s(c)}{H(c)}$$

Où H représente l'histogramme couleur et donc $SCR(c) \in [0,1]$. Une faible valeur de $SCR(c)$ veut dire qu'il y a une dispersion de la couleur c dans l'image, tandis que pour une couleur dominante homogène $SCR(c)$ sera proche de 1.

2.1.6 Descripteur par couleurs dominantes

L'utilisation d'histogrammes pour représenter la distribution de couleur présente quelques inconvénients : l'utilisation inutile de grande partie de l'espace mémoire. Toutes les classes ont la même taille. Les histogrammes à plusieurs dimensions sont ((creux)), c'est-à-dire que la majorité des cellules ne comptent aucun pixel . Enfin, du point de vue des mesures de similarité employées, les mesures traditionnelles effectuent uniquement une comparaison

cellule à cellule. Donc les signatures par couleurs dominantes ont été proposées pour résoudre ces différents problèmes.

Ce descripteur fournit une description compacte des couleurs représentatives dans une image. Il est défini par :

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

Où :

N : désigne le nombre de couleurs dominantes.

c_i : la valeur de la *i*ème couleur dominante.

p_i : un poids exprimant sa fréquence relative d'apparition dans l'image.

v_i : sa variance.

La cohérence spatiale s représente l'homogénéité spatiale des couleurs dominantes dans l'image.

Pour calculer des couleurs dominantes, il faut choisir un espace de couleurs uniforme au niveau de la perception humaine. Le descripteur par couleurs dominantes est invariant aux transformations de similarité.

Voici les deux descripteurs ci-dessous :

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

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

Le degré de similarité entre deux couleurs c_{1k} et c_{2l} , noté $a_{1k,2l}$, est défini par :

$$a_{1k,2l} = \begin{cases} 1 - \frac{\delta(c_{1k}, c_{2l})}{\delta_{max}} & \text{si } \delta(c_{1k}, c_{2l}) \leq T_d \\ 0 & \text{sinon} \end{cases}$$

Où $\delta(c_{1k}, c_{2l})$ désigne la distance euclidienne entre deux couleurs c_{1k} , c_{2l} est un seuil prédéfini.

2.2 Les descripteurs de textures

La texture est une caractéristique fondamentale des images car elle concerne un élément important de la vision humaine, c'est le second attribut visuel largement utilisé dans la recherche d'images par le contenu. Car elle permet de pallier certains problèmes posés par l'indexation par la couleur, notamment lorsque les distributions de couleur sont très proches.

Par définition la texture est comme la répétition d'un motif créant une image visuellement homogène. Plus précisément, la texture peut être vue comme un ensemble de pixels (niveaux de gris) spatialement agencés selon un certain nombre de relations spatiales, ainsi créant une région homogène.

L'étude de la texture des objets d'une image a plusieurs buts : obtenir des informations sur la nature d'un objet, segmenter l'image en régions homogènes, identifier la texture afin de la réduire à un ensemble de paramètres (compression d'images), recherche d'image par contenu, etc.

Parmi les approches les plus connues pour la caractérisation de texture, on peut citer : les méthodes statistiques, les méthodes fréquentielles et les méthodes géométriques. Elle repose sur l'analyse d'une fonction de densité spectrale dans un domaine fréquentiel.

2.2.1 Les méthodes statistiques

La majorité des méthodes d'analyse de texture sont basées sur des méthodes statistiques. Ce sont des méthodes qui étudient les relations entre un pixel et ses voisins, elles sont basées sur des évaluations quantitatives de la distribution de niveaux de gris.

Les méthodes statistiques étudient les relations entre un pixel et ses voisins et elles sont utilisées pour caractériser des structures fines, sans régularité apparente.

Les approches les plus utilisées sont : la méthode de la dépendance spatiale des niveaux de gris (SGLDM : Spatial Gray Level Dépendance Method) ou matrices de cooccurrences, caractéristiques de Tamura, la matrice de longueur de plages et la méthode de différence de niveau de gris (GLDM : Gray Level Différence Method).

2.2.1.1 Les matrices de longueur de plages (MLDP)

La MLDP est un descripteur statistique d'ordre supérieur de l'image, qui a surtout été utilisé pour la reconnaissance des textures. Ils sont une manière de représenter une texture par une estimation de la probabilité d'observer des plages des niveaux de gris constants dans une certaine direction. La MLDP regroupe, pour une région spécifiée de l'image, le nombre de plages ayant chaque longueur et valeur possibles. Elle regroupe donc le nombre de pixels successifs, sur une direction imposée, ayant le même couleur [25], [26]. Pour la description statistique de la texture, les valeurs de la MDLP sont combinées dans des descripteurs généralistes, qui relèvent des attributs statistiques [17].

– Le nombre de longueurs de plages :

$$SLP = \sum_{ng=0}^{Ng-1} \sum_{l=1}^L P_{\theta}(ng, l)$$

– La proportion de petites et grandes plages :

$$RF1 = \frac{1}{SLP} \sum_{ng=0}^{Ng-1} \sum_{l=1}^L \frac{P_{\theta}(ng, l)}{l^2}$$

$$RF2 = \frac{1}{SLP} \sum_{ng=0}^{Ng-1} \sum_{l=1}^L l^2 P_{\theta}(ng, l)$$

– L'hétérogénéité des niveaux de gris :

$$RF3 = \frac{1}{SLP} \sum_{ng=0}^{Ng-1} (\sum_{l=1}^L P_{\theta}(ng, l))^2$$

RF3 mesure la dispersion des plages entre les niveaux de gris.

– L'hétérogénéité des longueurs de plages : ce paramètre donne une information sur la dispersion des plages entre les longueurs.

$$RF4 = \frac{1}{SLP} \sum_{l=1}^L (\sum_{ng=0}^{Ng-1} P_{\theta}(ng, l))^2$$

– Le pourcentage des plages :

$$RF5 = SLP / R$$

$P_{\theta}(ng, l)$: indique le nombre de plages de pixels ayant le niveau de gris de longueur l.

Le paramètre θ indique la direction de la plage.

N_g : correspond au nombre de niveaux de gris dans l'image et L définit la longueur de la plage maximale.

K : Le nombre total des pixels de l'image.

2.2.1.2 Matrice de cooccurrence

Dans les années 70 Haralick et al [27] ont proposé une des premières méthodes de caractérisation de texture baptisée matrice de cooccurrence. C'est une méthode qui représente la référence en analyse statistique de texture d'une image. Les matrices de cooccurrence sont très riches en information de texture et servent souvent de méthode comparative pour les nouvelles approches. Elles font l'objet de plusieurs recherches. [28]

Cette méthode mesure la probabilité qu'un couple de niveaux de gris satisfaisant à une loi spatiale donnée apparaisse dans une image. Le niveau de gris des pixels de l'image dépend fortement des niveaux de gris des pixels voisins.

L'élément $p_{d,\theta}(i,j)$ de la matrice de cooccurrence présente la fréquence d'apparition des couples de niveaux de gris i et j pour les couples de pixels séparés par une distance d selon la direction θ . Le calcul de la matrice de cooccurrence nécessite le choix d'une distance et d'un angle de déplacement.

Pour obtenir de bons résultats, plusieurs chercheurs ont noté que la distance d'un pixel combinée avec des angles respectifs doit être de 0° , 45° , 90° et 135° .

C'est la solution que nous avons adoptée, ce qui nous donne à la fin quatre matrices de cooccurrence pour chaque image.

Les matrices sont ensuite normalisées par la formule suivante :

$$P_{d,\theta}(i,j) = \frac{p_{d,\theta}(i,j)}{M * N}$$

Pour une image de taille $M * N$.

Haralick [29] a proposé 14 attributs pouvant être extraits de ces matrices de cooccurrence, afin d'estimer la similarité entre ces dernières. Actuellement, seulement ces quatre caractéristiques sont largement utilisées: l'énergie, l'entropie, le contraste et le moment inverse de différence.

2.2.1.3 La méthode de différence de niveaux de gris (GLDM (Grey Level

Difference Method))

Cette méthode permet de calculer le nombre d'apparitions d'une différence de niveaux de gris donnée. Elle donne un aspect de la texture au sens de la différence de niveaux de gris entre les pixels. [28]

2.2.1.4 Caractéristique de Tamura

L'approche Tamura et al. [30] est intéressante pour la recherche d'images car elle décrit les textures possibles selon des concepts qui correspondent à la perception visuelle humaine. Les six propriétés visuelles des textures qui ont été proposées par ces auteurs sont : la grossièreté, le contraste, la direction, présence de lignes (linéarité), régularité et rugosité, Chacun deux et mesurer pour construire un vecteur de texture.

2.2.2 Les méthodes fréquentielles

L'une des méthodes de description de la texture les plus utilisées est la méthode fréquentielles. L'analyse fréquentielle d'images est une analyse qui est la plupart du temps déterminante de la texture étudiée. Elle repose sur l'analyse d'une fonction de densité spectrale dans un domaine fréquentiel. Parmi ses méthodes on cite, la transformée de Fourier, le filtre de Gabor, les ondelettes, etc. Ces dernières consistent à extraire l'énergie portée par le signal dans diverses bandes de fréquence.

2.2.2.1 Transformée de Fourier

La transformée de Fourier (TF) nous permet de passer d'une représentation de l'image dans le domaine spatial à sa représentation dans le domaine fréquentiel. Les résultats TF nous indiquent que nous pouvons extraire le contenu fréquentiel global de certains paramètres caractéristiques texturaux.

Pour une image I définie contenant N lignes et M colonnes, la représentation de la transformée de Fourier sera discrète. Elle est donnée par la formule suivante :

$$TFD_i(u, v) = \frac{1}{N * M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} I(i, j) \exp^{-f2\pi(\frac{ui}{N} + \frac{vj}{M})}$$

$$0 \leq u \leq N - 1$$

$$0 \leq v \leq M - 1$$

2.2.2.2 Transformée en ondelettes

Les ondelettes ont été introduites pour la première fois par Grossman et Morlet [31] comme un outil mathématique d'analyse des signaux sismiques, et depuis, elle a été largement utilisée dans des applications diverses en traitement d'images tels que le débruitage et la restauration [32] et la compression d'image. Une des transformées en ondelettes les plus couramment employées en analyse d'images est la transformée de Haar. Ils sont fréquemment employés en apprentissage pour avoir la description d'un objet (comme un visage ou une personne).

Cette transformée consiste à décomposer un signal en utilisant une famille de fonctions $\Psi_{m,n}(x)$ obtenues par la translation et la dilatation d'une fonction mère $\Psi(x)$, Les ondelettes suivantes sont générées à partir de cette fonction $\Psi(x)$.

$$\Psi_{m,n}(t) = \frac{1}{\sqrt{m}} \Psi\left(\frac{t - n}{m}\right) \quad m > 0, n \in R$$

Où :

$\Psi_{m,n}(x)$: les ondelettes.

M : est le facteur d'échelle (dilatation).

N : est le facteur de translation.

$\frac{1}{\sqrt{m}}$: est un facteur de normalisation à travers les différentes échelles.

Les coefficients de la transformée en ondelettes d'un signal $s(x)$, sont définis comme suit :

$$C_s(m, n) = \frac{1}{\sqrt{m}} \int_{-\infty}^{+\infty} s(x) \Psi\left(\frac{x-n}{m}\right) dx$$

2.2.2.3 Filtres spatiaux (Spatial domain filters)

Le filtrage spatiale est une méthode, en optique, permettant de sélectionner des composantes spatiales d'une image. Contrairement au filtrage temporel, on ne s'intéresse pas aux fréquences temporelles (les couleurs), mais aux fréquences spatiales (la forme de l'image). [33]

2.2.2.4 Filtres de Gabor

Le filtre de Gabor porte le nom de Dennis Gabor (1900-1979), un physicien anglais d'origine hongroise. Connu par son invention de l'holographie (qui lui vaudra le prix Nobel de physique en 1971), Il proposa la fonction 1-D de Gabor en 1946. Ce n'est qu'en 1985, que Jean Daugman propose la version 2D du filtre de Gabor. [34]

Les filtres de Gabor sont largement utilisés en indexation d'images pour la caractérisation de la texture et notamment pour la segmentation des images texturées et aussi dans le domaine de la reconnaissance faciale. Ils permettent une bonne résolution temporelle à haute fréquence et une bonne résolution harmonique sans grande précision temporelle à basse fréquence [29].

Il s'agit de la somme de deux fonctions sinusoïdales, l'une paire et réelle, l'autre impaire et imaginaire modulées par une courbe gaussienne, la fonction $g(x)$ est donc complexe. Dans le cas bidimensionnel, on va définir tout un ensemble de filtre gaussien séparable polairement. On subdivise ainsi le demi-plan fréquentiel ($f-f$) en N orientation et M bandes passantes. [20]

La fonction de transfert d'un filtre de Gabor bidimensionnel est défini par :

$$G(x, y) = \cos(W_x X + W_y Y) e^{-\frac{x^2+y^2}{\sigma^2}}$$

$$G(x, y) = \sin(W_x X + W_y Y) e^{-\frac{x^2+y^2}{\sigma^2}}$$

Les principales étapes d'extraction des caractéristiques des textures de l'image en utilisant le filtre de Gabor sont :

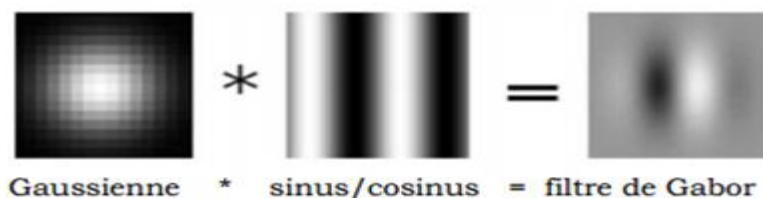


Figure 3.6 : Représentation d'un filtre de Gabor 2D.

2.3 Les descripteurs de formes

Au même titre que pour la texture, l'information de forme est complémentaire de celle de la couleur c'est l'une des principales caractéristiques visuelles importantes. Elle est une des caractéristiques de base pour décrire le contenu d'une image. L'extraction d'attribut géométrique a été le fer de lance de la recherche d'image par le contenu ces dernières années. Nous distinguons deux catégories de descripteurs de formes, les descripteurs basés régions : ces attributs sont robustes aux transformations géométriques comme la translation, la rotation et le changement d'échelle, ils font classiquement référence aux moments invariants et sont utilisés pour caractériser l'intégralité de la forme d'une région. Et les descripteurs basés frontières : ils font classiquement référence aux descripteurs de Fourier et porte sur une caractérisation des contours de la forme.

La forme est intéressante pour retrouver certains concepts qui ne peuvent l'être autrement. Prenons le mot ballon, par exemple. Il n'y a pas de couleurs qui puissent caractériser un ballon puisque cet objet peut être de toutes les couleurs, par contre, il y a une forme très caractéristique [20]. La forme a plusieurs avantages par rapport la texture.

Nous présentons dans ce qui suit quelques méthodes de description de la forme.

- **Les moments géométriques**

Les moments géométriques permettent de décrire une forme à l'aide des propriétés statistiques. Ils représentent les propriétés spatiales de la distribution des pixels dans l'image. Ils sont facilement calculés mais le temps de calcul de ces moments est très long en plus cette approche est très sensible au bruit et aux déformations. La formule générale des moments géométriques est donnée par la relation suivante:

$$m_{p,q} = \sum_{p=0}^m \sum_{q=0}^n x^p y^q f(x, y)$$

$p+q$ est l'ordre du moment. Le moment d'ordre 0 $m_{0,0}$ représente l'aire de la forme de l'objet. Les deux moments d'ordre 1 $m_{0,1}$ et $m_{1,0}$, associés au moment d'ordre 0 $m_{0,0}$ permettent de calculer le centre de gravité de l'objet. Les coordonnées de ce centre sont :

$$x_c = \frac{m_{1,0}}{m_{0,0}} \quad \text{et} \quad y_c = \frac{m_{0,1}}{m_{0,0}}$$

A partir de ces moments on peut calculer l'ellipse équivalente à l'objet. Afin de calculer les axes de l'ellipse, il faut ramener les moments d'ordre 2 au centre de gravité :

$$\begin{aligned} m_{2,0}^g &= m_{2,0} - m_{0,0}x_c^2 \\ m_{1,1}^g &= m_{1,1} - m_{0,0}x_c y_c \\ m_{0,2}^g &= m_{0,2} - m_{0,0}y_c^2 \end{aligned}$$

Puis on détermine l'angle d'inclinaison de l'ellipse α :

$$\alpha = \frac{1}{2} \arctan \frac{2m_{1,1}^g}{m_{2,0}^g - m_{0,2}^g}$$

- **Les moments invariants de Hu**

Ce sont des moments issus des moments d'inertie, nommés également moments de Hilbert Ils se construisent à partir du moment normé.

$$N_n = \frac{J^m K^n}{M_0^\alpha} \bar{M}_m$$

Ou : $\alpha = \frac{1}{2}(m + n) + 1$

À partir des moments géométriques, Hu a proposé sept invariants aux translations, rotations et changement d'échelle, appelés moments de Hu :

$$h_1 = N_{20} + N_{02}$$

$$h_2 = (N_{20} - N_{02})^2 + 4N_{11}^2$$

$$h_3 = (N_{30} - 3N_{12})^2 + (N_{03} - 3N_{21})^2$$

$$h_4 = (N_{30} + N_{12})^2 + (N_{03} + N_{21})^2$$

$$h_5 = (N_{30} - 3N_{12})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{03} + N_{21})^2] \\ + (3N_{21} - N_{03})(N_{03} + N_{21})[3(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2]$$

$$h_6 = (N_{20} - N_{02})[(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ + 4N_{11}(N_{30} + N_{12})(N_{03} + N_{21})$$

$$h_7 = (3N_{12} - N_{30})(N_{03} + N_{21})[3(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ + (3N_{21} - N_{03})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{03} + N_{21})^2]$$

Ces moments sont très utilisés dans la littérature pour la description des formes en vue d'une classification ou d'une indexation, mais sont assez sensibles aux bruits.

- **Les moments orthogonaux**

Contrairement aux moments géométriques qui sont définis par rapport à une base quelconque (x^p, y^q) les moments orthogonaux sont définis dans une base orthogonale, ce qui évite la redondance des informations portées par chacun des moments. Les moments orthogonaux les plus utilisés sont : les moments de Legendre et les moments de Zernike, dont nous donnons les définitions ci-dessous.

- **Les moments de Legendre**

Les moments de Legendre sont définis à partir des polynômes du même nom. Ils sont définis dans le carré unité $[-1,1]$ et $[-1,1]$, qui se lie à la normalisation de l'objet dont nous voulons calculer les moments. Le polynôme de Legendre d'ordre n est donné par :

$$\forall x \in [-1, 1], \forall n \in N, P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n}$$

Les polynômes de Legendre $(P_n(x))$ forment une base complète et orthogonale sur le domaine de définition $[-1,1]$:

$$\forall (x, y) \in [-1, 1]^2, \forall (m, n) \in N^2, \iint_{-1}^1 P_m(x) P_n(y) dx dy = \frac{2}{2m + 1} \delta_{mn}$$

δ_m représente la fonction de Kronecker. Les moments de Legendre d'ordre N sont donc donnés par :

$$\forall (x, y) \in [-1, 1]^2, \forall (p, q) \in N^2, N = p + q \\ L_{pq} = \frac{(2p + 1)(2q + 1)}{4} \iint_{-1}^1 P_p(x) P_q(y) f(x, y) dx dy$$

Où : $f(x,y)$ est le niveau de gris d'un pixel de l'image I sur laquelle on calcule le moment. Nous pouvons générer une infinité de moments de Legendre à partir de cette équation.

• **Les moments de Zernike**

La méthode des moments de Zernike est une autre méthode très utilisées pour extraire les caractéristiques globales d'une image en reconnaissance et en analyse d'images .Ces moments sont construits à partir de polynômes complexes et forment un ensemble orthogonal complet définie sur le disque unité. Grace à sa capacité de description et sa robustesse aux bruits et aux déformations Ce descripteur de forme a prouvé sa supériorité sur les autres fonctions de moments. Un moment de Zernike est un nombre complexe qui contient deux valeurs distinctes : le module et la phase.

Les moments orthogonaux de Zernike d'ordre p sont définis comme suit :

$$A_{mn} = \frac{m + 1}{\pi} \iint I(x, y) [V_{m,n}(x, y)] dx dy$$

Où m et n définissent l'ordre du moment et $I(x,y)$ le niveau de gris d'un pixel de l'image sur laquelle on calcule le moment Où $x^2 + y^2$.

Les polynômes de Zernike $V_{m,n}(x, y)$ sont exprimés en coordonnées polaires :

$$V_{mn}(r, \theta) = R_{mn}(r) e^{-jn\theta}$$

Où $R_{m,n}(r)$ est le polynôme radial orthogonal :

$$R_{mn}(r) = \sum_{s=0}^{\frac{m-|n|}{2}} (-1)^s \frac{(m-s)!}{s! \left(\frac{m-|n|}{2} - 2\right)! \left(\frac{m-|n|}{2} - s\right)!} r^{m-2s}$$

Avec $m = 0, 1, 2, 3, \dots, |n| \leq m$ et $|m - n|$ un entier pair.

Les polynômes de Zernike sont orthogonaux donc cette propriété d'orthogonalité supprime l'effet de redondance de l'information portée par chaque moment.

• **Descripteurs de Fourier**

Les Descripteurs de Fourier DFs sont parmi les descripteurs les plus connues pour les applications de reconnaissance de formes et de recherche d'images, grâce à leurs simplicités et leurs bonnes performances en terme de reconnaissance. De plus, ils permettent de décrire la forme des objets à différents niveaux de détail. La transformé est utilisée pour représenter les pixels d'une image selon les modèles périodiques, ils se caractérisent par leur fréquence et leur orientation. Ils permettent de décrire la forme de l'objet à différents niveaux de détails.

3. Mesure de similarité

La mesure de similarité quantifie la proximité des images dans l'espace des caractéristiques. Elle est souvent métrique, les images sont considérées ressemblantes si la distance est faible. La complexité de calcul d'une distance doit être raisonnable parce que dans un système CBIR cette tâche s'exécute en temps réel. D'autres paramètres entrent en jeu tel la dimension de l'espace caractéristique, la taille de la base, etc.... . La méthode naïve de recherche calcule la distance entre la requête et toutes les images de la base puis les ordonne selon leurs scores. Ceci par conséquent rend le temps de réponse proportionnel au nombre d'images ($O(N)$). Les

méthodes d'indexation du contenu permettent par ailleurs de réduire cette complexité comparée à la recherche séquentielle. Pour résumer, la mesure de similarité vérifie généralement les propriétés :

- **La perception** : Une faible distance dans l'espace caractéristique indique deux images semblables.
- **Le calcul** : La mesure de distance se calcule rapidement pour une faible latence.
- **La stabilité** : Le calcul de distance ne doit pas être affecté par une modification de taille de la base
- **La robustesse** : la mesure devra être robuste aux changements des conditions d'acquisition d'image.

3.1 Les méthodes de calcul

Ci-après les distances les plus utilisées pour comparer des images considérées comme vecteurs ou comme distributions statistiques.

3.1.1 Distance de Minkowski

La distance de Minkowski est une famille de distances vectorielles. Cette distance d'ordre r (ou D_r) est définie par :

$$D_r(X, Y) = \left(\sum_{i=1}^n |x_i - y_i|^r \right)^{1/r}$$

Où $r \geq 1$ est le facteur de Minkowski, et n la dimension de l'espace caractéristique.

Plus le paramètre r augmente plus la distance D_r aura tendance à favoriser les grandes différences entre coordonnées. La distance Euclidienne et la distance de Manhattan sont des cas particuliers de cette distance. Pour $r = 1$ on obtient la distance de Manhattan ou city block :

$$D_1(X, Y) = \left(\sum_{i=1}^n |x_i - y_i| \right)$$

Cette norme est plus appropriée pour mesurer la similarité entre les données multi-variées.

Pour $r = 2$, on obtient la distance Euclidienne:

$$D_2(X, Y) = \left(\sqrt{\sum_{i=1}^n |x_i - y_i|^2} \right)$$

La distance euclidienne est couramment utilisée dans des espaces à deux ou 3 dimensions, elle donne notamment de bons résultats s'il ensemble des données Présentes des classes compactes et isolées.

Pour $r = \infty$, on obtient la distance Chebyshev ou la distance maximum :

$$D_\infty(X, Y) = \lim_{r \rightarrow \infty} \sqrt[r]{\sum_{i=1}^n |x_i - y_i|^r} = \sup_i (|x_i - y_i|)$$

Cette distance est adaptée aux données de grandes dimensions.

3.1.2 Distance quadrature

La distance de Minkowski traite les éléments du vecteur de caractéristique d'une manière équitable.

La distance quadratique en revanche favorise les éléments les plus ressemblants. Sa forme générale est donné par :

$$D_Q = \sqrt{(f_1 - f_2)^T A (f_1 - f_2)} \quad \text{Ou} \quad A = [a_{ij}]$$

est la matrice de similarité. a_{ij} représente la distance entre deux éléments des vecteurs f_1 et f_2 .

Les propriétés de cette distance la rendraient proche de la perception humaine de la couleur, ce qui en fait une métrique attractive pour les systèmes de Recherche d'images couleur par le contenu.

3.1.3 Distance cosinus

La distance cosinus mesure la similarité entre X et Y en calculant le Cosinus de l'angle entre eux, ce qui détermine si deux vecteurs pointent dans la même direction. Elle est définie ainsi : [35]

$$S_{cos}(X, Y) = \cos\theta = \frac{X \cdot Y}{|X| \cdot |Y|} = \frac{\sum_{i=1}^n x_i X y_i}{\sqrt{\sum_{i=1}^n (x_i)^2} \times \sqrt{\sum_{i=1}^n (y_i)^2}}$$

3.1.4 Distance de Mahalanobis

Cette distance prend en compte la corrélation entre les distributions des classes. Elle est ainsi définie par : [15]

$$D_M = \sqrt{(f_1 - f_2)^T C^{-1} (f_1 - f_2)}$$

Où C est la matrice de covariance. Dans les cas où les dimensions des caractéristiques sont indépendantes, C ne comporte que des les variances et la distance de Mahalanobis se simplifie sous la forme :

$$D_M = \frac{\sum (f_1(i) - f_2(i))^2}{C_i}$$

Si C est la matrice identité, D_M est la distance euclidienne.

3.1.5 Distance de bhattacharya

La distance de Bhattacharya exploite la séparabilité entre deux distributions gaussiennes représentées par leur covariance Σ : [19]

$$D_B = \frac{1}{8} (\mu_1 - \mu_2)^T \Sigma^{-1} (\mu_1 - \mu_2) + \frac{1}{2} \ln \frac{\det(\Sigma)}{\sqrt{\det(\Sigma_1) \det(\Sigma_2)}}$$

Où $\Sigma = 0.5X(\Sigma_1 + \Sigma_2)$ La séparabilité entre classes est estimée par la distance des moyennes et des matrices de covariance de chaque classe.

3.1.6 Distance de Kullback Leiber(KL)

La divergence de Kullback Leibler es tissue de la théorie de l’information, la Divergence de Kullback Leibler se base sur l’entropie relative de la distribution X par rapport à Y, sa forme générale est exprimé ainsi:[15]

$$Dist_{kul}(X, Y) = \sum_{i=1}^n x_i \log \frac{x_i}{y_i}$$

3.1.7 Distance de Kolmogorov Smirnov

Cette distance est appliquée aux distributions cumulées $f^c(i)$:[15]

$$D_{ks} = \max_i |f_1^c(i) - f_2^c(i)|$$

3.1.8 Distance de Cramer Von Mises

La distance de Cramer Von Mises s’applique également sur des distributions cumulées, elle est définie par : [15]

$$D_{CVM} = \sum_i (f_1^c(i) - f_2^c(i))^2$$

3.1.9 Intersection d’histogramme

Cette mesure est l’une des premières distances utilisée dans la recherche d’image par le Contenu. Elle a été proposée par Swain et Ballard mesurant la partie commune entre deux histogrammes. Etant donné deux histogrammes $h1$ et $h2$. [15]

$$D_{Intersec} = \frac{\sum_i \min(h_1(i), h_2(i))}{\sum_i h_2(i)}$$

Deux images présentant une intersection normalisée d’histogrammes proche de 1 sont considérées comme similaires. Cette mesure n’est pas une métrique parce que non symétrique. Cependant il en existe des versions symétriques telles que celle proposée par Smith.

3.1.10 Earth Mover Distance(EMD)

EMD consiste à minimiser le coût de transformation d’une distribution en une autre sous certaines contraintes de déplacement des classes de descripteurs. EMD requiert une optimisation linéaire. [15]

$$D_{EMD} = \frac{\sum_{ij} g_{ij} d_{ij}}{\sum_{ij} g_{ij}}$$

Ou d_{ij} représente la disimilarité entre deux indices (i, j) et g_{ij} est le flot optimal entre deux distributions dont le coût total est :

$$D_{EMD} + \sum_{ij} g_{ij} d_{ij}$$

Le coût est minimisé sous les b :

$$g_{ij} \geq 0, \forall i, j$$

$$\sum_i g_{ij} \leq f_2(j), \forall j$$

$$\sum_j g_{ij} \leq f_1(i), \forall i$$

$$\sum_i \sum_j g_{ij} = \min(f_1(i), f_2(j))$$

EMD prétend également mimer la vision humaine.

3.1.11 Divergence de Jeffrey (JD)

La divergence Jeffrey est la forme symétrique de la divergence de Kullback- Leibler, elle est définie comme suit :

$$Dist_{DJ}(X, Y) = \sum_{i=1}^n \left[x_i \log \frac{x_i}{m_i} + y_i \log \frac{y_i}{m_i} \right]$$

Ou :

$$m_i = \frac{(x_i + y_i)}{2}$$

4. Conclusion

Le choix des descripteurs pour un système de recherche d'images par contenu est important, dans le sens où, ce choix influe sur les résultats attendus. Cependant, d'une part il n'y a pas d'attributs universels, et d'autre part le choix des descripteurs dépend fortement de la base d'image à utiliser et des connaissances à priori qu'on peut avoir sur la base.

CHAPTRE 04 :



Le Travail Effectué

1. Introduction

Après avoir étudié le domaine de CBIR, l'implémentation d'une application d'un système de recherche d'images devient une nécessité afin d'avoir une vue plus claire. Ce chapitre donne une vue plus détaillée sur les techniques utilisées dans notre travail.

2. Méthode utilisée

Dans ce chapitre, nous allons parler de la manière dont nous pouvons implémenter les règles que nous avons étudiées dans les chapitres précédents, nous allons proposer une approche qui permet de récupérer les images similaires de la base de données en utilisant une image comme requête. Le but de ce travail est de récupérer les images de façon rapide et moins complexe. L'idée de cette approche est d'extraire la texture et la couleur qui caractérisent l'image.

3. Architecture

Pour simplifier les choses, nous avons préparé une architecture qui résume les étapes et les méthodes que nous avons effectuées lors de notre mise en œuvre du logiciel, voir la (figure 4.1)

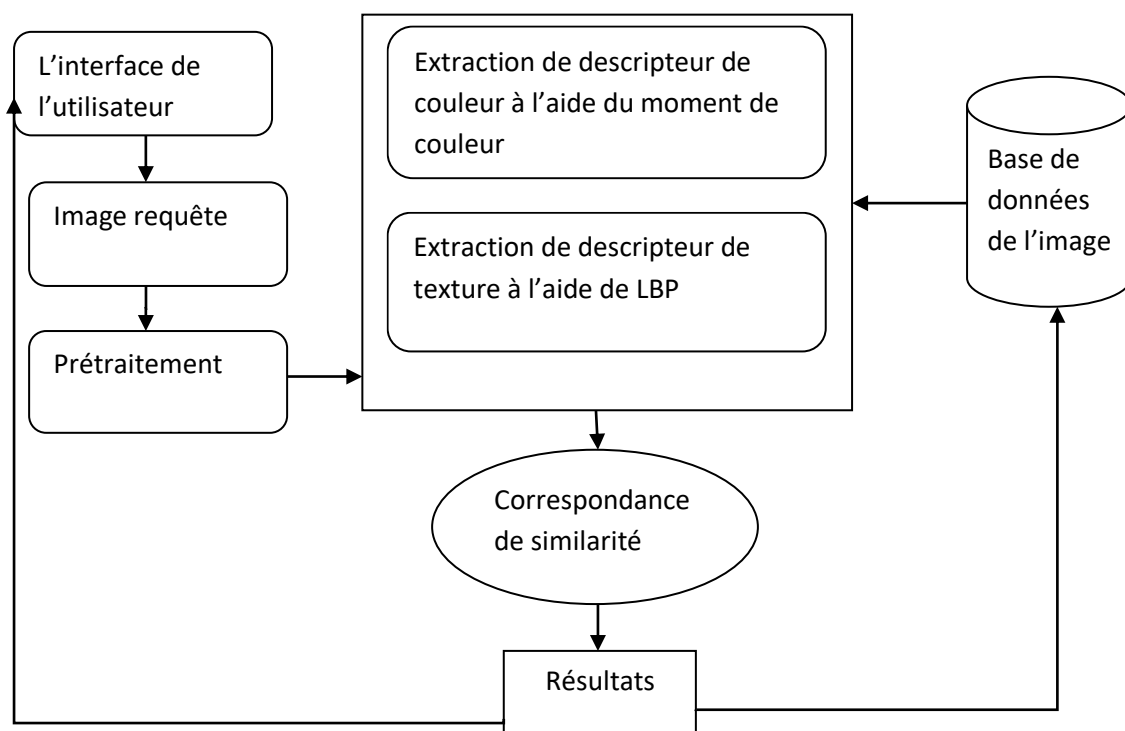


Figure 4.1 : Architecture de travail effectué

Premièrement, l'image est prétraitée afin que le descripteur peut être extrait. Deuxièmement, le prétraitement de la texture et la couleur sont combiné pour extraire le vecteur descripteur. Enfin, les résultats sont récupérés après l'extraction de ce dernier et la correspondance de similarité. La description de l'architecture est abordée ci-dessous :

3.1 Le prétraitement

C'est un processus qui est appliqué sur l'image avant l'extraction de descripteur, il comprend l'amélioration, la segmentation et la conversion de l'image, etc. L'amélioration est utilisé pour améliorer le contraste de l'image. La segmentation, est utilisé pour créer des segments de l'image en fonction de la couleur ou de segmenter en bloc. La conversion est le processus de conversion de l'image d'un espace colorimétrique à un autre.

3.2 L'extraction de descripteur

C'est un élément important de la récupération d'image basée sur le contenu, il est basé sur l'extraction des descripteurs d'image. Les techniques utilisées sont définies ci-dessous :

- **local binary pattern (LBP)**

Ce descripteur a été mentionné pour la première fois en 1993 pour mesurer le contraste local d'une image mais réellement popularisé trois ans plus tard par Ojala et al [38] pour analyser les textures. Le principe général est de comparer le niveau de luminance d'un pixel avec les niveaux de ses voisins. Cela rend donc compte d'une information relative à des motifs réguliers dans l'image, autrement dit une texture. Selon l'échelle du voisinage utilisé, certaines zones d'intérêt tels des coins ou des bords peuvent être détectées par ce descripteur. [36]

Le descripteur LBP n'est pas coûteux en calcul. Pour créer une représentation LBP nous devons d'abord convertir l'image couleur (espace colorimétrique RVB) en image en niveaux de gris. Ensuite, nous calculons la valeur LBP. En utilisant ces valeurs, nous obtenons une image LBP et calculons l'histogramme des codes LBP calculées comme indiqué dans la (figure 4.2).

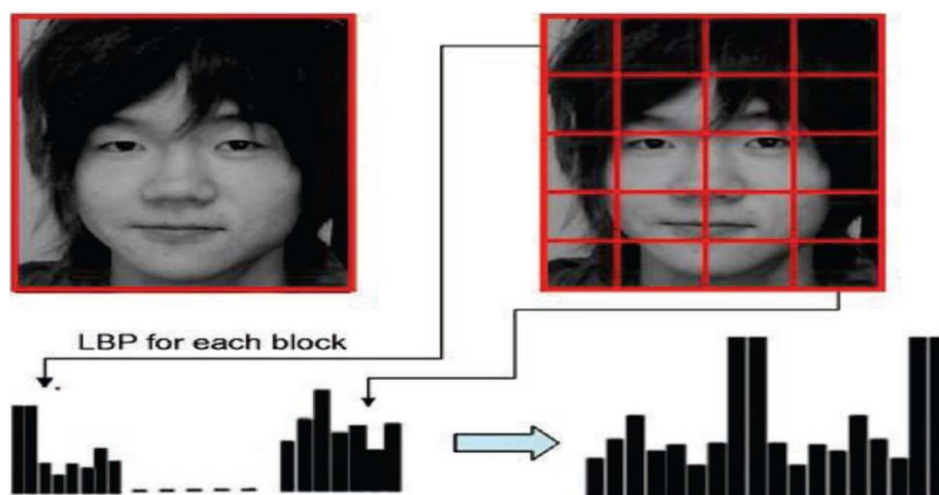


Figure 4.2 : Descripteur LBP

Ci-dessous (le tableau 1) montre la valeur d'un sous-bloc de taille 3X3 en niveaux de gris.

180	168	149
175	150	120
157	100	133

Tableau 1 : sous-bloc de l'image de taille 3X3

Bp ₀	Bp ₁	Bp ₂
Bp ₇	Cp	Bp ₃
Bp ₆	Bp ₅	Bp ₄

Tableau 2 : Position du bit de sous-bloc de l'image

(Le tableau 2) montre la position de bit différente avec le pixel central sur le sous-bloc. Toutes les valeurs du pixel sont utilisées pour comparer avec le pixel central 'cp' pour calculer les codes LBP.

Maintenant nous allons comparer tous les positions des bits bp₀, bp₁, bp₂... etc, avec le pixel central cp en utilisant la formule suivante :

$$bp_i = f(x) = \begin{cases} 1: \sum_{i=0}^7 bp_i \geq cp \\ 0: \sum_{i=0}^7 bp_i < cp \end{cases}$$

Après avoir appliqué la formule ci-dessus, nous obtenons (le tableau 3) et (tableau 4) en tant que modèle binaire.

1	1	0
1	P	0
1	0	0

Tableau 3 : sous-bloc 3X3 montrant la valeur binaire

1	2	4
128	Cp	8
64	32	16

Tableau 4 : sous-bloc 3X3 valeur pondérée de la position du bit

Enfin, nous devons trouver la chaîne binaire qui représente un motif en utilisant la formule suivante:

$$Cp = \sum_{i=0}^7 2^i * bp_i$$

B^i est un bit dans (le tableau 3). Maintenant, chaque position de bit (de tableau 3) et de (tableau 4) sont multipliées à l'aide de la formule ci-dessus et nous obtenons une valeur de fonctionnalité LBP qui agit comme une signature de la texture dans (le tableau 5).

1	2	0
128		0
64	0	0

Tableau 5 : Sous-bloc 3X3 de la valeur du descripteur LBP

Ensuite, nous ajoutons toute la position du bit et obtenons la valeur de $cp = 1 + 2 + 128 + 64 = 195$.

L'histogramme du motif LBP de tout pixel central cp est alors créé pour décrire la texture de l'image.

$$H_{ij} = \sum_{X,Y} [\{f_i(X,Y) = i\} \{ (X,Y) \in R_j \}]$$

L'histogramme extrait représente la texture locale de l'image.

- **les moments de couleur**

Nous avons choisi les moments de couleur comme une technique de base pour la description de la couleur dans notre application. Comme on a déjà mentionnée dans le chapitre 3, Les moments de couleur est une technique utilisée pour extraire la couleur de l'image, ils sont utilisés pour différencier les images en fonction de la couleur de l'image et pour vérifier la couleur similitude entre les images.

4. Mesure de similarité

La mesure de distance est la mesure d'une certaine distance entre deux images qui sont comparées [37]. On a utilisé la distance euclidienne comme mesure de similarité, elle est donnée par la formule suivante :

$$D(F_{V_q}, F_{V_{DB_{im}}}) = \{ \sum_{i=0}^{L-1} (F_{V_q}(i) - F_{V_{DB_{im}}}(i))^2 \}^{1/2}$$

Ou :

$$F_{V_q}(i) = (F_{V_q}(0), F_{V_q}(1), \dots, F_{V_q}(L-1))$$

$F_{V_q}(i)$ est le vecteur descripteur de l'image requête.

Where Ed (IrsK) > Th

L'algorithme démarre lorsque l'utilisateur fournit l'image de requête via l'interface utilisateur.

6. Bases d'images utilisé

Nous évaluons les performances de notre système avec des bases d'images connues par la communauté qui travaille dans ces technologies. Dans notre travail on a utilisé la base de Wang, cette base contient 1000 images naturelle en couleur répartie en 10 classes : Afrique, Plage, Bâtiments, Autobus, Dinosaures, Fleurs, Eléphants, Chevaux, Alimentation et Montagnes, chaque classe contient 100 images et chaque image est de taille 384x256.

7. Résultat expérimental

Ici, on va comparer les deux descripteurs, les moments de couleurs et le LBP(figure 4.4) montre les résultats en utilisant seulement les moments de couleur sur une image requête (figure 4.3).

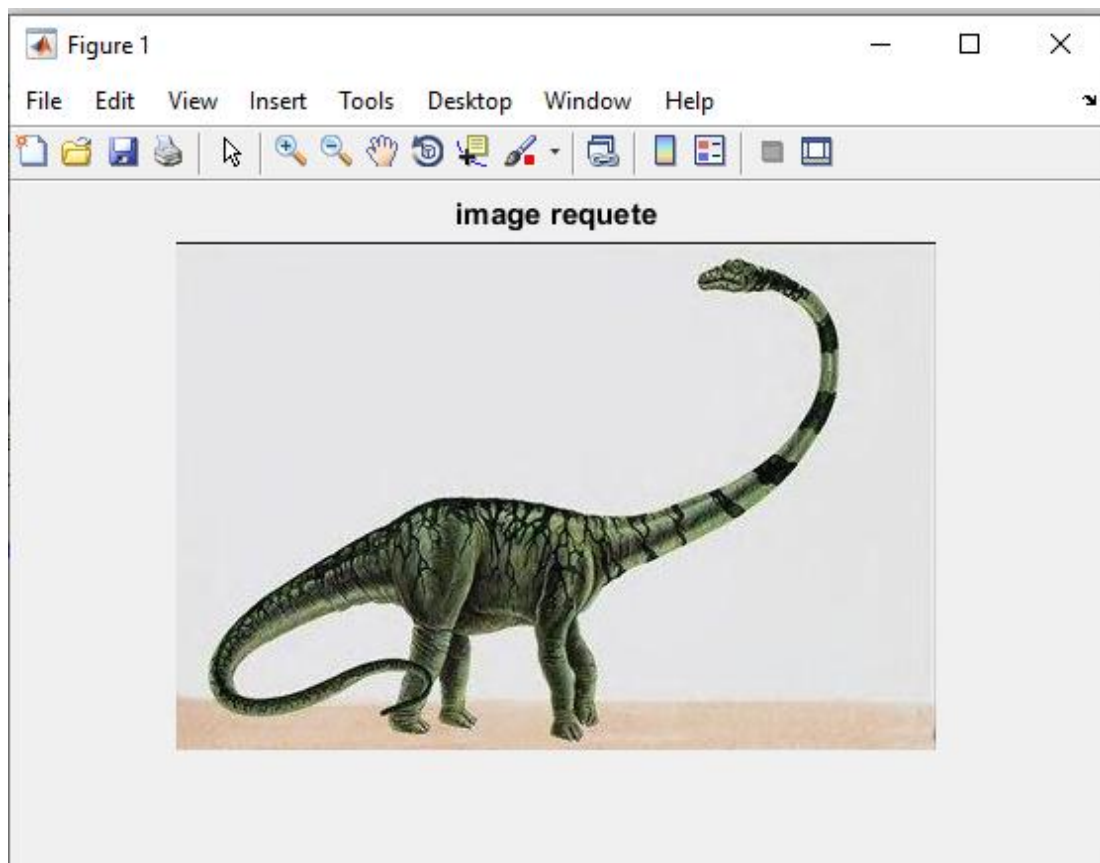


Figure 4.3 : Image requête



Figure 4.4 : Résultat de moment de couleur

L'image de requête illustrée à (la figure 4.3) est utilisée encore pour récupérer le résultat à l'aide de LBP, les résultats à l'aide de cette méthode sont illustré à la (figure 4.5).



Figure 4.5 : Résultat de LBP

Le but c'était de voir est ce qu'on peut utiliser qu'un seul descripteur pour la récupération des images similaire, donc par les résultats obtenues dans (la figure 4.4) et (figure 4.5) nous voyons quelles sont pas précise, alors nous avons combiné deux techniques sur l'image requête, (figure 4.6).

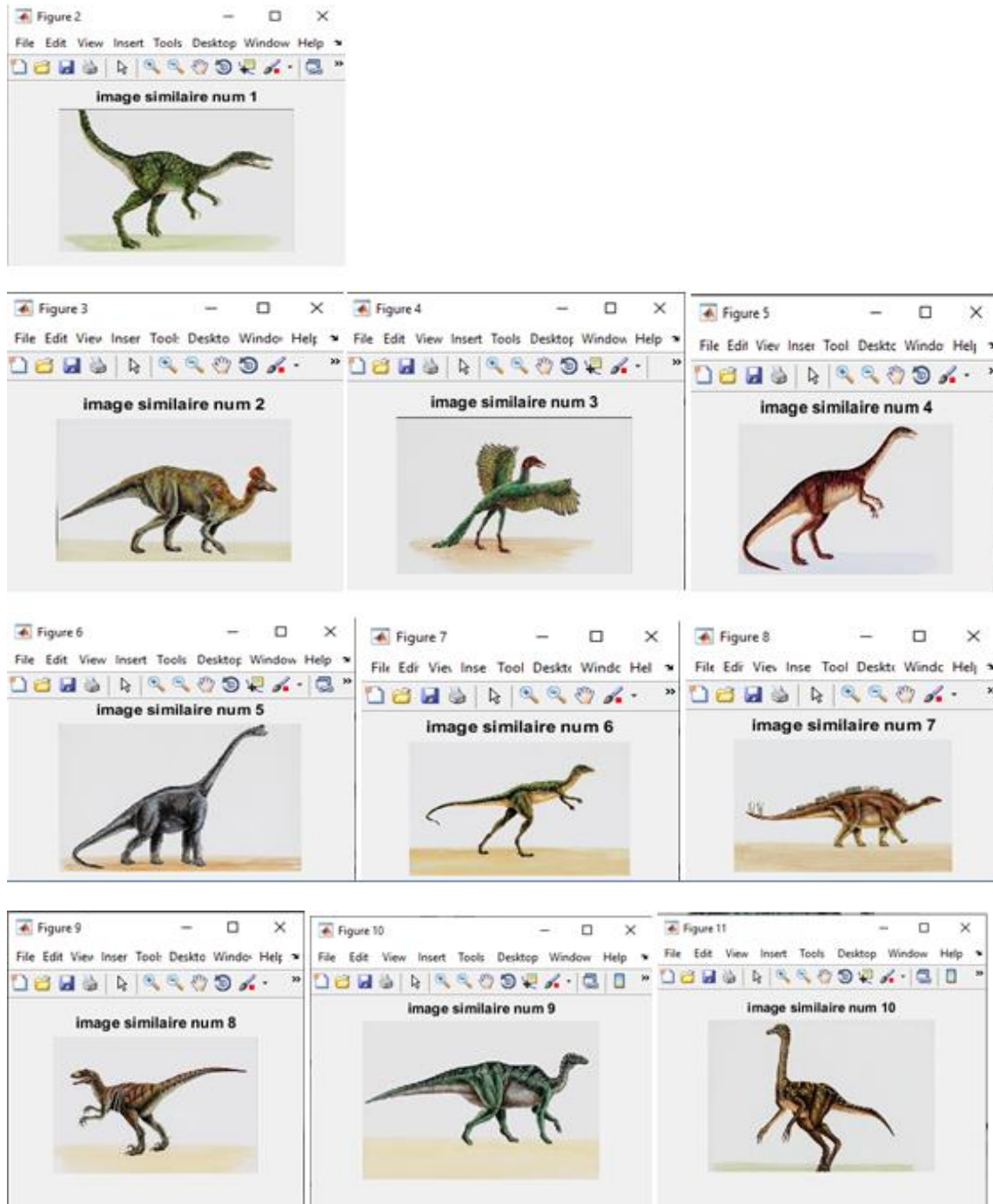


Figure 4.6 : Résultat de vecteur descripteur

En conclusion, et d'après les résultats obtenues, en utilisant les deux descripteurs de couleur et texture pour récupérer les images similaire de la base d'images, donne des bons résultats plutôt que d'utiliser un seul descripteur, cette méthode est plus précise, et moins complexe par rapport à d'autre.

8. Conclusion

Nous avons présenté dans ce chapitre notre contribution en essayant de répondre aux besoins des utilisateurs liés au domaine de la recherche d'images par le contenu. Il s'agit d'une approche intégrée utilisée pour extraire la couleur et la texture des images. La première étape a été de définir des descripteurs adaptés aux images testées. Pour extraire la fonction de couleur, les moments de couleur sont utilisés comme descripteur de couleur. Pour extraire la texture sur des images naturels, LBP est utilisé comme descripteur de texture, qui est généralement utilisé pour la reconnaissance faciale. L'utilisation d'un seul descripteur n'est pas suffisante et pas précise, contrairement à l'utilisation d'un vecteur descripteur qui donne des résultats satisfaisante.

Conclusion générale

Dans ce mémoire, nous avons abordé le problème de la recherche et l'indexation d'images, plus précisément nous nous sommes intéressées sur la recherche et l'indexation d'images basée sur le contenu visuel. Notre choix a été motivé par la quantité d'images disponibles aujourd'hui et qui augmentent de jour en jour. Il s'agit d'extraire, à partir des images, une certaine information pertinente de couleur, de texture et de forme. Ces informations permettant une recherche efficace, pour atteindre cet objectif. Un système d'indexation et de recherche d'images par le contenu a été proposé, ce système fonctionne avec des descripteurs visuels qui sont extraits d'une image telle que les caractéristiques de l'histogramme RVB et niveau de gris.

Notre objectif était de mettre en œuvre un maximum de traitements d'images en niveaux de gris possibles, en se basant sur la programmation sous environnement Matlab sans l'utilisation des fonctions prédéfinies de ce dernier.

Du point de vue de recherche, nous avons utilisé le principe de la recherche par similarité. Cette dernière prend la forme d'une image exemple. Durant ce travail nous avons constaté deux points essentiels et cruciaux. Le premier, est que la couleur est une caractéristique discriminante d'une image, mais l'utilisation de la couleur tout seul dans un CBIR ne suffit pas, il faut rajouter d'autres descripteurs de texture et de forme pour la performance d'un CBIR.

Le domaine de recherche d'images par le contenu est un domaine très riche et exhibe une variance dans les techniques utilisées. Il n'existe pas une loi qui impose le choix d'une technique particulière pour l'extraction des images en se basant sur le contenu.

Les références :

- [1] : Diane .L, INTRODUCTION AU TRAITEMENT D'IMAGES, Vuibert, paris, 2004.
- [2] : Gérard .M, Images numériques et Formats graphiques, Paris, décembre 2000.
- [3] : Virgine c, « Quels sont les 9 formats différents pour une image », Webdesign, 6 novembre 2012.
- [4] : Henri Maitre, Le traitement des images, La voisier, 2003.
- [5] : Zerougui. A et Sari. N, Traitement d'images monochromes, mémoire de master, université de Oum el bouaghi, 2017.
- [6] : Boukhlouf. D, chapitre3 : généralités sur traitement d'image, Université de Biskra, 2005.
- [7] : Kaidi. D, Classification non supervisée de pixels d'images couleur par analyse d'histogrammes tridimensionnels, mémoire de master, université de Tizi- Ouzou, 2017.
- [8] : Melliani. M, segmentation d'image par coopération régions- Contours, mémoire de magister en informatique, école national supérieur d'informatique, Alger, 2012.
- [9] : Sandeli. M, traitement d'images par des approches bio-inspirées application à la segmentation d'images, mémoire de magister, université constantine 2, 2014.
- [10] : Houassine. C, segmentation d'images par une approche biomimétique Hybride, mémoire de magister, université de m'hamed bougara-Boumerdès, 2012.
- [11] : Belarouci .S, Benmokhtar .S, méthode coopérative pour la segmentation d'images IRM cérébrales basée sur les techniques FCM et level Set, 2012.
- [12] : Canny. J, A, computational approach to edge detection, IEEE trans. on P.atern analysis and machine intelligence, vol. 8, p 679-698, novembre 1986.
- [13] : Deriche. R, Using Canny's criteria to derive a recursively implemented optimal edge detector, international journal of computer vision, pp. 167-187, 1987.
- [14] : M. Kass, A. Witkin And D. Terzopoulos , snakes : active contour models, cinternational journal of computer vision, p321–331, 1987.
- [15] : Kamel. H, Recherche d'images par le contenu, Thèse de Doctorat, Université Mentouri de Constantine, 2010
- [16] : Wang. J, Champs Markoviens multi-échelles applications à la

- Segmentation d'images texturées et à la fusion multi-film. Thèse de Doctorat, Université Paris XI, Centre d'Orsay France, 1994.
- [17] : Dr. Alain.B, indexation et recherche d'image par le contenu, institut Polytechnique de HANOI, mémoire de master, 2005.
- [18] : BEDOUHENE.S, Recherche d'images par le contenu, mémoire de magister, université Moloud Mammeri, Tizi-ouzou ,2011.
- [19] : MERABET.N, MAHLIA.M, Recherche d'images par le contenu, mémoire de master, Université Abou Bakr Belkaid-Tlemcen, 2011.
- [20] : Berradj. K, Diaf .D, Recherche d'image par le contenu, mémoire de master, université Mouloud Mammeri de Tizi-Ouzou, 2018.
- [22] : Adjemout. O, Reconnaissance automatique de formes à partir des Paramètres morphologiques, de couleur et de texture : application au tri des graines de semences, mémoire de magister, UMMTO, 2005.
- [23] : M.A. Stricker and M. Orengo. Similarity of color images. In SPIE, Storage and Retrieval for image Video Databases, pages 381-392, 1995.
- [24]: Meskaldji. M, Extraction et traitement de l'information : Un prototype d'un système de recherche d'images couleurs par le contenu magistère, Université Mentouri de Constantine, 2009.
- [25] : M. M. Galloway, "Texture analysis using gray level run lengths", In graphical Models and image Processing. Vol. 4. Pp.172-179, 1975.
- [26] : J. P. Cocquerez, Philipp S. (coord.), Analyse d'images : filtrage et segmentation, Masson, Paris, 1995.
- [27] : R.M. Haralick, K. Shanmugan, I. Dinstein, Textural features for image Classification. IEEE Trans. Syst. Man. Cybern. SMC-3(6): 610–621. 1973.
- [28] :LAIB .H, BENSGHIER .A, Indexation Et Recherche D'image Fixe Basé Sure Le Contenu, mémoire de master, université de MOHAMMED SEDDIK BENYAHIA JIJEL, 2020
- [29] : D. Mercier, R. Séguier, Utilisation des STANN en audio : illustration reconnaissance de chiffre, Journée Valgo 2001.
- [30] : H. Tamura, S. Mori, and T. Yamawaki, Texture features corresponding to visual perception, IEEE Trans. On Systems, Man, and Cybernetics, vol. Smc-8, No. 6, Juin 1978.
- [31] :A. Grossmann and J. Morlet. Decomposition of Hardy functions into Square integrable wavelets of constant shape. SIAM J. Math. Anal., 15(4) : 723–736, 1984.
- [32] : A. Khare and U. S. Tiwary . A New Method for Deblurring and Denoising

Of Medical Images using Complex Wavelet Transform. Proceedings of the 2005 IEEE.

- [34] : M. AISSOU Bahaeddine, Analyse des images satellitaires à haute résolution, mémoire de magister, UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE D'ORAN Mohamed BOUDIAF, 2014.
- [35] : Benkrama . S, Indexation et recherche d'images par contenu dans les grands corpus d'images, Thèse de Doctorat, université de Oran, 2017
- [37] : Jau-Ling Shi h and Ling-Hwei Chen, "Color Image Retrieval Based on Primitives of Color Moments", IEE E Proceedings- Vision Image and Signal Processing, volume 149 Issue 6, pp 370-376, 2002.
- [38] : T.Ojala,M. Piet ik in en,D.Harwood, « A comparative study of texture measures with classification based on featured distributions », pattem Recognition, vol.29,No.1,pp.51-59,1996.

Webliographie :

- [21] : https://fr.wikipedia.org/wiki/Teinte_Saturation_Valeur.
- [33] : https://fr.wikipedia.org/wiki/Filtrage_spatial
- [36] : https://fr.wikipedia.org/wiki/Motif_binaire_local