

République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université de Guelma

Faculté des Mathématiques, d'Informatique et des Sciences de la matière

Mémoire de Master



Département d'Informatique

Spécialité : Ingénierie des Médias

Construction automatique de bases d'apprentissage

Présenté par : Zedadra Amina

Bouchelaghem Zahra

Sous la direction de :

Pr. Seridi Hamid

Juin 2011

Remerciements

C'est avec fierté que nous apportons ce modeste travail à tous ceux qui nous ont gratifiés de leur soutien et de leur confiance.

Louanges à Dieu, qui nous a donné vie et santé pour le parachèvement de ce travail.

Nous tenons à exprimer notre profonde gratitude en premier lieu à notre co-encadreur Monsieur **Toualbiya Ilyes**, Maître assistant à l'université de Guelma.

Nous tenons à remercier très vivement et chaleureusement notre encadreur Monsieur **Seridi Hamid**, professeur à l'université de Guelma, pour avoir bien voulu accepter d'être encadreur de ce travail. Nous le remercions pour sa disponibilité, ses encouragements, son soutien moral et sa sympathie dont il nous fait preuve le long de ce travail. Nous apprécions la confiance qu'il nous a témoignée et les conseils avisés qu'il nous a prodigués.

Nous remercions aussi et plus chaleureusement **Mlle. Zedadra Nawel** et **Mlle. Zedadra Saida** pour leurs patiences et leurs soutiens indéfectibles.

Nous ne serons sans doute pas arrivé jusqu'ici sans avoir suivi les cours d'enseignants exceptionnels qui ont participé à notre formation, je pense en particulier aux enseignants du département informatique de Guelma.

Nous remercions nos familles, qui durant nos études, nous ont toujours donné la possibilité de faire ce que nous voulions et qui ont toujours cru en nous.

Enfin, nous remercions tous ceux qui ont soutenu et aider de près ou de loin pour la réalisation de ce travail.

Zedadra Amina & Bouchelaghem Zahra

Résumé

L'objectif de notre travail est de réaliser un système qui permet la construction automatique de bases d'apprentissage « **ACLB** : Automatic Construction of Learning Base ».

Le système « **ACLB** » consiste à chercher automatiquement du web un ensemble d'images selon une requête posée par l'utilisateur. Un dictionnaire dans le moteur de recherche est utilisé pour enrichir automatiquement les différentes requêtes posées. Afin de restreindre l'ensemble des images collectées à un ensemble d'images pertinentes, des méthodes de recherche d'images par contenu ont été utilisées. Aussi des méthodes de détection d'objets pour filtrer les résultats en gardant que les images qui contiennent l'objet sélectionné par l'utilisateur. Nous offrons une étude comparative entre les résultats des différentes méthodes implémentées. Cette évaluation facilite à l'utilisateur le choix de la méthode à appliquer pour construire sa base d'apprentissage.

À la fin de la collecte d'images pertinents, nous procédons à l'extraction d'attributs caractéristiques pour limiter la quantité d'informations utilisée dans l'apprentissage et améliorer la qualité d'apprentissage en ne gardant que les attributs caractéristiques discriminants et éliminant ceux redondants et le bruit. Ces attributs caractéristiques sont utilisés pour construire la base d'apprentissage.

Mots clés : Base d'apprentissage, recherche automatique, recherche par contenu d'image, enrichissement de requêtes, RDF et détection d'objets.

Sommaire

Sommaire	1
Liste des figures	6
Liste des tableaux	8
Introduction général	9

Chapitre I : Recherche d'image dans le web

1. Introduction	11
2. Recherche d'images dans le web (point de vue utilisateur et système)	11
2.1. Intention de l'utilisateur	13
1) Navigateur	13
2) Internaute	13
3) Chercheur	14
2.2. Champ d'application des données	14
1) Collection personnelle	14
2) Collection d'un domaine spécifique.....	14
3) Collection d'une entreprise	14
4) Archive.....	14
5) Web	14
2.3. Modalités et traitement des requêtes	15
2.4. La visualisation	16
1) Ordre de pertinence	16
2) Ordre chronologique	16
3) Cluster	16
4) Hiérarchique.....	17
5) Composite	17
3. Types de recherche d'image	17
3.1. Recherche par texte	17
3.2. Recherche par image	18
2.2.1. Les avantages de cette méthode par rapport à la méthode basée sur le texte	18
2.2.2. Les difficultés rencontrées avec ces méthodes.....	18
4. Quelques applications de la recherche d'images sur le web	19

4.1.	Préservation de la propriété intellectuelle	19
4.2.	Filtrage de contenu pour adultes	19
4.3.	Application de la loi et la prévention du crime	20
4.4.	Voyage et tourisme	20
4.5.	Éducation et formation	20
4.6.	Loisirs	20
5.	Quelques techniques associées à la recherche d'images sur le web	20
5.1.	Collecte de données	20
5.2.	Conversion d'image	20
5.3.	Validité d'image	21
5.4.	Descripteurs d'image	21
5.5.	Similarité et correspondance	21
5.6.	Spécification de requête	21
5.7.	Récupération et raffinement	22
5.8.	Évaluation de la performance	22
6.	Présentation de LYCOS	21
6.1.	La formulation syntaxique des requêtes	22
6.2.	La fonction de filtrage	23
6.3.	Présentation des résultats	23
6.4.	La démarche suivie pour la recherche sur LYCOS.....	23
7.	Conclusion	24

Chapitre II : Recherche d'images pas le contenu

1.	Introduction	25
2.	Historique de la recherche d'images par le contenu	25
3.	Généralités de la recherche d'images	26
4.	Les méthodes de recherche par le contenu	26
4.1.	Les caractéristiques de bas niveau	26
4.1.1.	La couleur	26
4.1.2.	La texture.....	29
4.1.3.	La forme	30
4.2.	Les méthodes sémantiques	31
4.2.1.	Les travaux réalisés	32

5. Conclusion	38
Chapitre III : Reconnaissance des formes et détection d'objets	
1) Introduction	39
Partie I: La reconnaissance des formes :	
I.1 Définition des RDF.....	39
I.2 Le rôle de reconnaissance des formes	40
I.3 Schéma général d'un système de Reconnaissance de Formes	40
I.3.1. L'espace de mesure	41
I.3.2. L'espace de représentation	41
I.3.3. L'espace de décision ou d'interprétation	41
I.4 Méthode RDF	42
1) Méthode Bayesienne	43
2) Classifieur linéaire	43
3) Réseau de neurones	43
4) SVM : Support Vector Machine	43
I.5 Champs d'Applications	44
1) Reconnaissance de la Parole	44
2) Reconnaissance de l'écriture	44
3) La vision	44
Partie II: Détection d'objets	
II.1. Définition	44
II.2. Les algorithmes de reconnaissance d'objet.....	45
1) Méthodes généralistes	45
2) Méthodes de recherche basées sur la transformation	45
3) Méthodes guidées par des modèles géométriques	45
4) Méthodes basées sur un modèle 3D	45
5) Méthodes de construction de contour flexible et contour active.....	45
6) Méthodes basées sur des descripteurs	45
II.3. Domaines d'application	45
1) Localisation de la position d'objet (Position measurement)	46
2) Inspection	46
3) Trie (Sorting).....	46
4) Comptage	46
5) Détection d'objet	46

6) Catégorisation de scène	46
7) Recherche d'images (Image Retrieval)	46
II.4. Exigence et contraintes	47
1) Temps d'évaluation (Evaluation Time)	47
2) Précision	47
3) Fiabilité de reconnaissance	47
4) Invariance	47
5) Luminosité.....	47
6) Taille de la forme	47
7) Rotation.....	47
8) Chevauchement avec l'arrière plan	48
9) Occultation d'une partie de l'objet	48
10) Changement de l'angle de capture	48
II.5. Classification des méthodes de reconnaissance	48
1) Représentation de l'objet	49
2) Domaine des données de l'objet	49
3) Variations possibles de l'objet	49
4) Qualité des données de l'image	50
5) Le processus de reconnaissance	50
6) Domaine des éléments de données utilisées dans la mesure de correspondance	50
2) Conclusion	50

Chapitre IV : Conception et implémentation

1. Introduction	51
2. Conception	51
2.1. Objectifs du système	51
2.2. Architecture du système.....	52
2.2.1. Web	52
2.2.2. Récupération des images à partir du web	53
2.2.3. Enrichissement des requêtes.....	54
2.2.4. Recherche d'image par contenu	54
2.2.4.1. Recherche d'image en utilisant la couleur.....	54

2.2.4.2.	Recherche d'image en utilisant la texture	55
2.2.5.	Filtrage à base d'objet choisi (méthode de corrélation)	57
2.2.5.1.	Corrélation	57
2.2.5.2.	Corrélation (variante Sobel)	59
3.	Implémentation de système	62
3.1.	Aspect matériel	62
3.2.	Environnement de développement	62
3.2.1.	Présentation de java.....	62
3.2.2.	Pourquoi java ?.....	63
3.2.3.	Présentation d'Eclipse	63
3.2.4.	Présentation de l'image dans java	63
3.3.	La réalisation de l'application	64
3.4.	Présentation de l'application	66
3.4.1.	Fenêtre principale	66
3.4.2.	Chargement d'image requête	67
3.4.3.	Méthode de recherche par la couleur	68
3.4.4.	Méthode de recherche par la texture	69
3.4.5.	Combinaison des deux méthodes (couleur et la texture)	69
3.4.6.	La méthode de la corrélation	70
3.4.7.	La méthode de la corrélation (variante Sobel)	71
3.4.8.	Construction de la base.....	71
4.	Etude comparative	72
4.1.	Comparaison des méthodes de recherche d'image par contenu.....	72
4.2.	Comparaison entre les méthodes de détection d'objet et les méthodes de recherche d'image par contenu	74
5.	Conclusion	75
	Conclusion générale et perspectives	76
	Bibliographie	77
	Annexe	79

Liste des figures

Chapitre I : Recherche d'images dans le web

Figure I.1: Recherche des images de point de vue utilisateur	12
Figure I.2: Recherche des images de point de vue système.....	13

Chapitre II : Recherche d'images par le contenu

Figure II.1 : Extraction des descripteurs pour la recherche d'images.....	26
Figure II. 2 : Présentation des couleurs dans l'espace RGB.....	28
Figure II.3 : Présentation des couleurs dans l'espace TSV.....	28
Figure II.4 : Recherche d'image en se basant sur la texture de l'image .	30
Figure II.5 : Découpage en grille de l'image	30
Figure II.6 : Découpage par la segmentation division-fusion de l'image	31
Figure II.7 : Architecture du modèle basé sur l'intelligence artificielle et la reconnaissance des formes.....	32
Figure II.8 : Résultat du modèle sur l'intelligence artificielle et la reconnaissance des formes	34
Figure II.9 : Exemple du processus de retour de pertinence	35
Figure II.10 : Un exemple de retour de pertinence dans le CBIR.....	36
Figure II.11: Exemple de construction d'une ontologie.....	37
Figure II.12 : Cas d'utilisation d'une ontologie	37
Figure II.13 : Exemple d'une ontologie « Annotation d'images biomédicales et raisonnement »	38

Chapitre III: La reconnaissance des formes et détection d'objets

Figure III.1 : Schéma général d'un système de reconnaissance de forme	40
---	----

Chapitre IV: Conception et implémentation

Figure IV.1 : Étapes de construction de la base d'apprentissage.....	52
Figure IV.2 : Présentation du site web utilisée par notre système « LYCOS ».....	53
Figure IV.3 : Présentation des résultats de recherche sur « LYCOS ».....	53
Figure IV.4 : Présentation des recherches liées à la requête sur « LYCOS ».....	54
Figure IV.5 : Les étapes de recherche d'image en se basant sur l'histogramme..	55
Figure IV.6 : Les étapes de recherche d'image en se basant sur la texture.....	56
Figure IV.7 : Présentation des distances et des directions..	56
Figure IV.8 : Exemple de matrices de cooccurrences construites à partir d'une image 4x5 composées de huit niveaux de gris.....	57
Figure IV.9 : Les étapes de la méthode de corrélation.....	58

Figure IV.10 : Exemple d'application de la méthode de corrélation..	59
Figure IV.11 : Les étapes de la méthode de la corrélation(variante Sobel)..	59
Figure IV.12 : Les étapes de la détection des contours.....	60
Figure IV.13 : Méthode de détection des contours en utilisant le filtre de Sobel..	61
Figure IV.14 : Environnement Eclipse.....	63
Figure IV.15 : Représentation d'un pixel couleur.....	64
Figure IV.16 : Exemple extrait du code.....	65
Figure IV.17 : Démarrage de l'application.	66
Figure IV.18 : Fenêtre principale de « ACLB ».....	67
Figure IV.19 : Chargement de l'image requête.....	68
Figure IV.20 : Méthode de recherche par la couleur.....	68
Figure IV.21 : Méthode de recherche par la texture.	69
Figure IV.22: Combinaison des deux méthodes (couleur et la texture).....	70
Figure IV.23 : La méthode de la corrélation..	70
Figure IV.24 : La méthode de la corrélation (variante Sobel)..	71
Figure IV.25 : La base d'apprentissage.....	72

Liste des tableaux

Tableau IV.1 : Comparaison des méthodes de recherche d'images par contenu	72
Tableau IV.2 : Distances en fonction de la taille d'image.....	73
Tableau IV.3 : Distances en fonction de la position d'image.....	73
Tableau IV.4 : Valeurs de rappel et précision avec tous les descripteurs implémentés	73
Tableau IV.5 : Comparaison des méthodes de détection d'objet	74
Tableau IV.6 : Valeurs de rappel et précision avec toutes les méthodes implémentées	74

Introduction générale

Introduction Générale

Deux décennies auparavant, Tim Berners-Lee mit à disposition du monde les premières pages web, elles peuvent contenir différents types de données et d'informations : textes, images, vidéo, audio, etc., et font référence à d'autres pages web. Aujourd'hui, le web est devenu le plus grand gisement de données, source inépuisable d'informations que chacun de nous utilise pour se documenter et apprendre. Techniquement, Internet est la plateforme physique qui relie des millions d'ordinateurs éparpillés partout dans le monde, tandis que web est le réseau virtuel qui relie de grandes quantités d'informations et données stockées sur ces ordinateurs.

Les utilisateurs effectuent, le plus souvent, la recherche d'informations sur le web en faisant recours aux moteurs de recherche. Ces derniers ne sont pas toujours parfaits, Si vous utilisez Google par exemple, vous tapez une requête, il retourne beaucoup d'information dont peu n'ont aucun rapport avec la requête.

Le système ACLB est un système conçu pour améliorer les résultats de recherche d'image dans le web, tout en récupérant les résultats d'une requête d'utilisateur et les enregistrent sur disque puis on applique des méthodes de recherche par contenu et aussi des méthodes de détection d'objet, afin de garder que les résultats (images) pertinents répondant exactement à la requête d'utilisateur. A chaque fois qu'on applique une de ces dernières méthodes, nous procédons à l'extraction d'attributs caractéristiques (à savoir la moyenne, la variance, l'inertie, l'entropie, l'énergie et l'homogénéité). Nous avons utilisé quatre critères (précision, rappel, nombre des images pertinentes par rapport à ceux retournées par le système et le temps d'évaluation) pour évaluer les différentes méthodes de recherche ainsi implémentées. Cette évaluation facilite à l'utilisateur le choix de la méthode à appliquer pour construire sa base d'apprentissage.

Notre mémoire s'articule autour de quatre chapitres :

Le premier chapitre est consacré à la recherche d'image dans le web et le choix du moteur de recherche. Dans ce chapitre nous présentons les méthodes ainsi que les enjeux et les techniques associés à cette recherche, puis nous choisissons le moteur de recherche approprié.

Le deuxième chapitre porte sur la recherche des images par contenu, dans lequel nous présentons quelques généralités sur les différentes méthodes de recherche, tels que les méthodes de recherche par caractéristiques de bas niveau ou par contenu sémantique.

Le troisième chapitre est composé de deux parties, l'une sur la reconnaissance des formes et l'autre sur la détection d'objet.

Dans le dernier chapitre nous présentons la conception et l'implémentation de du système de construction automatique de bases d'apprentissage « **ACLB** : **A**utomatic **C**onstruction of **L**earning **B**ase».

Enfin, nous terminons ce travail par une conclusion générale et des perspectives.

Chapitre I : Recherche d'images dans le web

- 1- Introduction
- 2- Recherche d'images sur le web (point de vue utilisateur et système)
- 3- Type de recherche
- 4- Quelques applications de la recherche d'images sur le web
- 5- Quelques techniques associées à la recherche d'images sur le web
- 6- Présentation de LYCOS
- 7- Conclusion

Chapitre I

Recherche d'images dans le web

1. Introduction

Avec le développement de la technologie Internet, le nombre des images disponible sur le Web augmente et il continue à croître chaque jour. Les utilisateurs récupèrent généralement l'information en utilisant les moteurs de recherche [1]. Lors de la recherche sur le web, le moteur de recherche permet de fournir à l'utilisateur un ensemble de documents dont peu sont pertinents.

En plus du texte, du son et de la vidéo, l'image constitue une autre modalité contenue dans le web. Les moteurs de recherches d'images utilisent principalement des informations textuelles, tels que le titre de la page web, le nom de l'image, le texte adjacent, pour tenter de comprendre le sens de l'image. Cependant, le texte d'une page web n'est pas toujours en rapport avec le contenu visuel de l'image. De plus, l'utilisateur préfère souvent exprimer son besoin d'information à l'aide de quelques mots-clés, mais il est difficile de trouver des liens avec l'information visuelle contenue dans les images [2].

Nous présenterons dans ce chapitre quelques notions de la recherche d'images sur le web (point de vue utilisateur et système), les types de recherche, les enjeux et les techniques associées à la recherche des images. Enfin nous terminerons par une présentation concise du moteur de recherche LYCOS.

2. Recherche d'images sur le web (point de vue utilisateur et système)

Nous expliquons dans cette partie les contextes de recherche d'images dans le monde réel, en explorant les attentes des utilisateurs, l'effort de recherche que ceux-ci doivent faire pour réussir leur recherche, les contraintes et les exigences que doivent respecter les systèmes de recherche pour répondre aux attentes des utilisateurs.

L'idée d'un moteur de recherche d'images puissant, capable de servir toutes les catégories d'utilisateurs est loin d'être résolue, il faut comprendre et caractériser les interactions utilisateur-système lors de la recherche d'images, à la fois de point de vue utilisateur et système [3].

La **figure I.1** présente toutes les interactions utilisateur-système lors de la recherche d'images. Premièrement, du point de vue utilisateur quand il commence à faire une recherche d'images, la procédure adéquate doit respecter les points suivants :

- ✓ La clarté de l'information voulue.
- ✓ Le choix de l'outil convenable pour obtenir l'information voulue.
- ✓ La formulation de la requête pour obtenir l'information voulue.

Deuxièmement, du point de vue système de récupération d'images, les facteurs pris en considération sont les suivants :

- ✓ Comment l'utilisateur souhaite que les résultats soient présentés .
- ✓ D'où vient le désir utilisateur de rechercher.
- ✓ La nature de la requête formulée par l'utilisateur pour interagir avec le système.

Ces facteurs avec leurs respectives valeurs possibles forment les axes de la **figure I.2**. Ils forment un repère 3D constitué des trois axes suivant : le champ d'application des données, la nature de l'utilisateur et les moyens d'interaction.

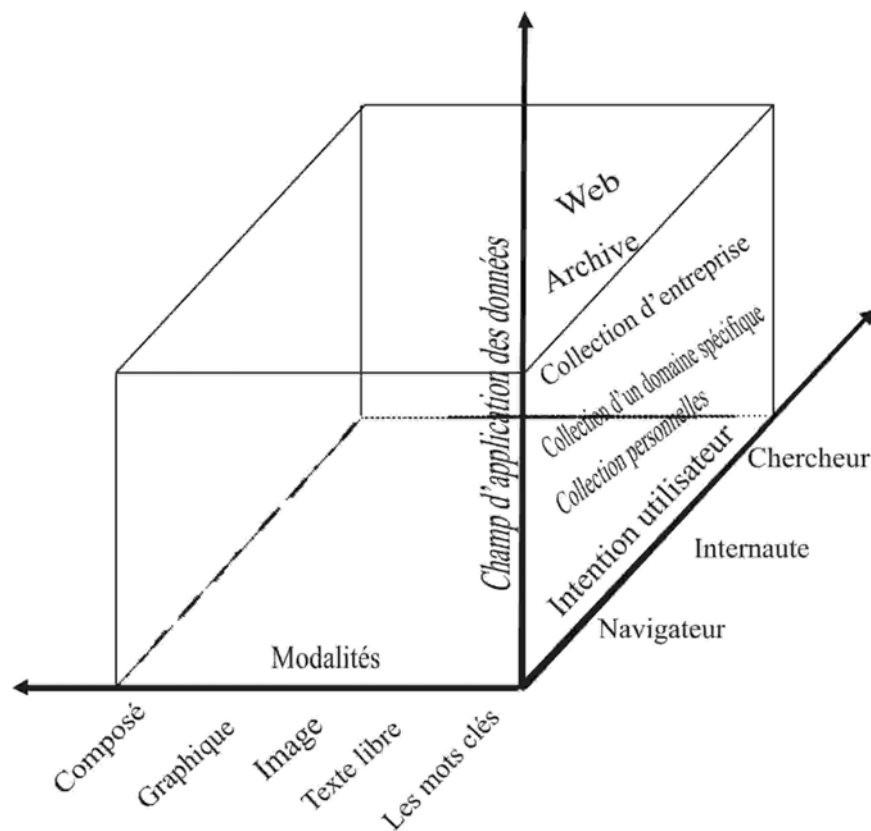


Figure I.1 : Recherche d'images du point de vu utilisateur.

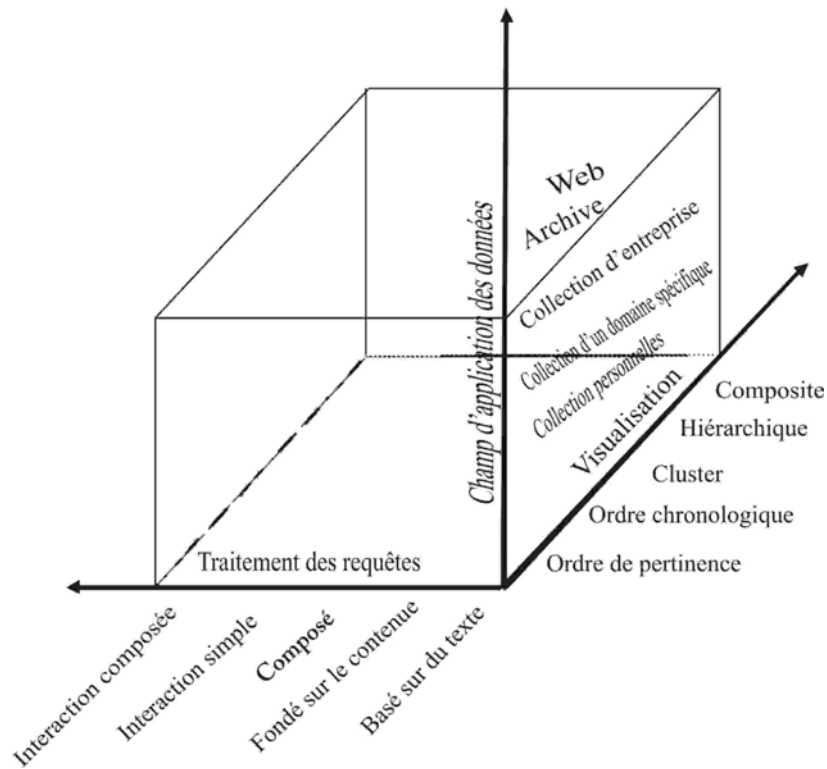


Figure 1.2 : Recherche d'images de point de vue système.

2.1. Intention de l'utilisateur

La recherche est basée sur une classification fondée sur l'intention de l'utilisateur. Quand les utilisateurs font la recherche d'images, leur intention ou la clarté de ce qu'ils veulent trouver peuvent varier. Ces derniers jouent un rôle important dans la suite des interactions entre les utilisateurs et le système de recherche. Les concepteurs doivent prendre en considération ce changement d'interaction dans la conception de leurs systèmes de recherche. Un utilisateur est caractérisé comme suit :

1) Navigateur

Il s'agit d'utilisateurs qui errent sur le web sans un objectif final clair. Une session de navigation consisterait en une série de recherches indépendantes. Un navigateur typique qui sautait à travers plusieurs thèmes au cours d'une session de recherche. Ses requêtes seraient incohérentes et diversifiées dans le sujet [3].

2) Internaute

Un internaute est un utilisateur qui sait ce qu'il cherche, ses actions peuvent être exploratoires au début, mais deviennent de plus en plus claires après quelques tentatives de recherches pour exprimer clairement ce qu'il attend du système [3].

3) Chercheur

Il s'agit d'un utilisateur qui sait parfaitement ce qu'il veut du système de recherche. Ses sessions de recherche sont courtes, cohérentes et finissent par atteindre le plus souvent son objectif [3].

2.2. Champ d'application des données

La nature et le champ d'application des données jouent un rôle clé dans la conception d'un système de recherche d'images. La classification des données de recherche se résume dans les catégories suivantes.

1) Collection personnelle

Il s'agit d'une collection très homogène généralement de petite taille, accessible principalement à son propriétaire, et généralement stocké sur un support de stockage local.

2) Collection d'un domaine spécifique

Il s'agit d'une collection homogène donnant accès à des utilisateurs contrôlés avec des objectifs très précis. La collection peut être grande et hébergé sur un système de stockage distribué. Des exemples d'une telle collection sont les bases de données biomédicales et les images satellite.

3) Collection d'une entreprise

C'est un ensemble hétérogène d'images accessibles aux utilisateurs au sein de l'intranet d'une organisation. Les photos peuvent être stockées dans de nombreux endroits. L'accès peut être uniforme ou non uniforme, en fonction du design de l'intranet.

4) Archive

Ces derniers sont généralement d'intérêt historique et contiennent de grands volumes des données homogènes structurées ou semi-structurée relatives à des sujets spécifiques. Les archives peuvent être accessibles à la plupart des gens sur l'internet, avec un certain contrôle de l'usage. L'archive est habituellement stockée dans de multiples supports de stockage de grande capacité.

5) Web

Les photos du Web sont accessibles à presque tout le monde avec une connexion internet. Les moteurs de recherche d'images actuels comme Google et Lycos font régulièrement la mise à jour de leurs bases de données locales pour refléter la nature dynamique du Web. La

collection d'images est semi-structurée, non homogène, et massive dans le volume, et est généralement stockée dans de grands disques [3].

2.3 Modalités et traitement des requêtes

Dans le domaine de la recherche d'images, un paramètre important pour mesurer le niveau d'interaction utilisateur-système est la complexité des requêtes posées au système. Du point de vue utilisateur, cela se traduit par les différentes modalités qu'il peut utiliser pour interroger un système. Une description des différentes modalités de requêtes, leurs caractéristiques, et le système nécessaire sont donnés comme suit [3] :

- **Les mots-clés** : il s'agit d'une recherche dans laquelle l'utilisateur pose une simple requête sous la forme d'un mot ou bigame. C'est actuellement le moyen le plus populaire pour rechercher des images.
- **Texte libre** : c'est là où l'utilisateur pose une phrase complexe, une question, ou un récit de ce qu'il veut du système.
- **Image** : ici, l'utilisateur souhaite rechercher une image semblable à une image requête. L'utilisation d'une image exemple est peut être la façon la plus représentative de l'interrogation d'un système de CBIR¹ en l'absence de métadonnées fiables.
- **Graphiques** : il s'agit d'une image dessinée à la main ou générée par ordinateur, fournie comme une requête au système.
- **Composé** : ce sont des méthodes qui impliquent l'utilisation d'un ou plusieurs des modalités susmentionnées pour l'interrogation d'un système.

Les modalités de requête ci-dessus requièrent différentes méthodes de traitement et/ou de soutien pour l'interaction utilisateur. Le traitement devient plus complexe lorsque des requêtes visuelles et/ou les interactions des utilisateurs sont impliquées [3].

Le traitement des requêtes dans le système est :

- **Basé sur du texte** : Le traitement des requêtes basé sur du texte revient habituellement à effectuer une ou plusieurs recherches basées sur des mots clés simples pour récupérer les photos correspondantes aux mots clés utilisés. Le traitement d'un texte libre pourrait comporter toute une analyse, une transformation

¹ CBIR : Content Based Image Retrieval

et une compréhension de la requête. En outre, un certain traitement automatique du langage naturel peut également être impliqué.

- **Fondé sur le contenu** : Le traitement des requêtes basées sur le contenu se trouve au cœur de tous les systèmes CBIR. Traitement de la requête (image ou graphique) implique l'extraction de caractéristiques visuelles et/ou la segmentation et la recherche dans l'espace des caractéristiques visuelles des images similaires.
- **Composé** : Le traitement composé peut impliquer à la fois le contenu et le traitement à base de texte dans des proportions variables.
- **Interaction simple** : L'interaction avec l'utilisateur en utilisant une modalité unique doit être soutenue par le système. Un exemple est le système de recherche des images basé sur un retour de pertinence.
- **Interaction composée** : L'utilisateur peut interagir en utilisant plus d'une modalité (par exemple, texte et images). C'est peut-être la forme la plus avancée du traitement des requêtes qui doit être effectuée par un système de récupération d'image.

2.4 La visualisation

La présentation des résultats de recherche est peut être l'un des facteurs les plus importants dans l'acceptation et la popularité d'un système de recherche d'images. Nous caractérisons la visualisation pour la recherche d'image comme suit :

1) Ordre de pertinence

Le moyen le plus populaire de présenter les résultats de recherche est la relevance ordonnée, adopté par Google et Yahoo pour leurs moteurs de recherche d'images. Les résultats sont commandés par une mesure numérique de la pertinence à la requête.

2) Ordre chronologique

Le résultat de la recherche est affiché dans un ordre chronologique plutôt que par pertinence.

3) Cluster

Dans ce cas, l'affichage se fait en regroupant les images selon leurs contenus visuels ou les métadonnées utilisées dans leur indexation. Cette démarche, intuitive et souhaitable dans la présentation des résultats, sert également à améliorer les performances de récupération.

4) Hiérarchique

Si les métadonnées associées aux images peuvent être disposés en ordre de l'arbre, ce dernier est très utile pour la visualisation hiérarchique des résultats de recherche, très souhaitable pour les archives, en particulier à des fins éducatives.

5) Composite

La composition consiste à mélanger deux ou plusieurs des formes précédentes de visualisation. Le regroupement hiérarchique et la visualisation de graphes conceptuels sont des exemples de visualisations composites.

3. Types de recherche d'image

3.1. Recherche par texte

La recherche des images basées sur le texte est actuellement la plus utilisée. Elle n'utilise pas les données des images lors du processus d'indexation et de recherche, mais des éléments textuels extrapicturaux.

- ✓ Dans le cas d'images présentes dans des collections, l'indexation s'effectue manuellement. L'indexeur crée une notice qui décrit l'image grâce à des éléments textuels. Plusieurs difficultés sont imposées lors de ce processus :
 - Diversité des niveaux de description et d'interprétation des images.
 - Polysémie.
 - Manque de reproductibilité inter-indexeur.
 - Coût en temps.
 - Problème de l'adéquation de l'indexation avec les aspirations des utilisateurs de la collection d'images, évolution du langage d'indexation avec le temps [4].
- ✓ Dans le cas des images présentes sur les sites Internet, l'indexation par les outils de recherche d'images est totalement automatique, sans intervention humaine. Le repérage des images s'effectue grâce à leur extension et le texte indexé avec l'image (qui est extrait du code HTML de la page) [4]. Il s'agit le plus souvent du nom de l'image (beaucoup d'images portent des noms génériques comme img001.jpg), texte de l'attribut ALT de la balise HTML IMG (pas toujours renseignée, mais qui est censée fournir une description de l'image) ou bien du texte proche de l'image (pas toujours en rapport avec l'image).

3.2. Recherche par image

Les méthodes de recherche par image, dénommées *recherche d'images basée sur le contenu*, sont par comparaison beaucoup plus récentes (le terme CBIR est apparu pour la première fois en 1992).

Avec ce type de méthodes, la phase d'indexation des documents n'implique à aucun moment l'utilisation du langage. Elle consiste à extraire un certain nombre (x par exemple) de paramètres d'analyse d'images (paramètres de couleur, texture, intensité ou forme), qui vont caractériser de manière discriminante les images présentes dans la base de données. Chaque image est alors identifiée non pas par des éléments textuels, mais par les valeurs de ses x paramètres [4].

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

3.2.1. Les avantages de cette méthode par rapport à la méthode basée sur le texte

Les avantages sont de plusieurs types :

- ✓ De par leur mode de fonctionnement, elles n'obligent pas à conceptualiser la recherche avec des mots et s'affranchissent donc des difficultés liées au langage rencontrées avec les premières.
- ✓ Ces méthodes permettent en outre de répondre à des questionnements ne faisant pas appel à un cadre sémantique (en permettant par exemple de rechercher des images contenant certaines gammes de couleurs ou d'effectuer une recherche à partir d'un exemple d'image existante ou créée pour l'occasion).

3.2.2. Les difficultés rencontrées avec ces méthodes

- × Les difficultés rencontrées avec ces méthodes sont principalement liées au fait qu'il n'est pas aisé de prendre compte de la perception visuelle de l'image par peu de paramètres.
- × Il peut ainsi exister un certain déphasage entre ce que la méthode de recherche par l'image considère comme des images similaires d'un point de vue paramétrique et ce que l'œil de

l'utilisateur considère comme des images similaires d'un point de vue visuel (c'est particulièrement le cas en ce qui concerne la texture des images).

- × De plus, ces méthodes sont relativement complexes à mettre au point et nécessitent un certain savoir-faire en traitement et analyse d'images.

Pour maximiser les performances de ces méthodes, il faut en effet trouver des paramètres discriminants pour différencier les images et, si possible, en petit nombre pour faciliter les traitements ultérieurs. Si cela est souvent possible lorsque les images sont de même type et contiennent des invariants (image de visage, d'empreintes digitales, etc.), il est en revanche très difficile de trouver une combinaison de paramètres discriminants pour des images dont le contenu varie grandement (ce qui est évidemment le cas pour les images présentes sur Internet).

Enfin, si les méthodes de recherche d'images par le texte permettent de répondre à des interrogations de bas niveau (recherche d'objets, personnes, etc.) ou de niveau plus élevé (recherche d'attributs abstraits des images), les méthodes de recherche par l'image permettent en revanche de répondre presque uniquement à des interrogations de bas niveau concernant les caractéristiques primitives des images (forme, intensité, texture, couleurs) ou, au mieux, la reconnaissance d'objets d'un type donné [4].

4. Quelques applications de la recherche d'images sur le web

4.1. Préservation de la propriété intellectuelle

La protection du droit d'auteur de l'image sur le web est essentielle, car les copies non autorisées des images peuvent être facilement transférées sur Internet. Une technique existante pour la protection du droit d'image est *watermarking* [5], cette technique permet d'ajouter pour chaque image l'identité du créateur ou bien du distributeur. À cause de la vulnérabilité de cette méthode, la recherche des images par contenu a fourni une bonne alternative, c'est le cas où l'image originale devient la requête et récupère la version suspect à partir du web.

4.2. Filtrage de contenu pour adultes

Les solutions anciennes comprennent sites Web libres (mais eux ne couvrent qu'une petite partie du Web), et un logiciel de filtrage tel que NetNanny, Cyber Patrol et Cyber Sitter. Récemment, des chercheurs ont exploré l'utilisation des techniques de vision par ordinateur

pour identifier automatiquement les images pornographiques (basée sur la texture de la peau, couleur de la peau et la forme du membre).

4.3. Application de la loi et la prévention du crime

En dehors du web, la recherche d'images basée sur le contenu a été utilisée pour la reconnaissance d'empreintes digitales, reconnaissance faciale, correspondance d'ADN, des empreintes de semelles de chaussures et de systèmes de surveillance. La recherche d'images sur le web peut être appliquée dans d'autres domaines d'application de la loi. Beaucoup de criminels utilisent le net comme un moyen pour promouvoir leurs biens et/ou de services illicites, tels que la vente des armes et des drogues. Beaucoup de ces sites contiennent peu de texte et plus d'informations visuelles. Dans ce contexte, la recherche à base de texte est inefficace à les localiser [5].

4.4. Voyage et tourisme

Pour rechercher des cartes des pays, l'architecture, les costumes traditionnels, les sites touristiques, etc.

4.5. Éducation et formation

La recherche d'image sur le web est très utile pour les étudiants (information, illustration, etc.) et les enseignants qui préparent des cours. Beaucoup d'images utiles sont disponibles sur le Web pour le téléchargement libre. C'est l'application la plus importante de la recherche d'images à partir du Web.

4.6. Loisirs

Images blague, bandes dessinées, des extraits de film, des clips vidéo.

5. Quelques techniques associées à la recherche d'images sur le web

5.1. Collecte de données

Pour rassembler les images, les crawlers identifient les éléments pertinents dans les sites web (par exemple, les images elles-mêmes, le texte, les adresses URL), extraient ses images pour les indexer et les enregistrer sur un support de stockage, puis suivent les URLs pour les examiner de manière récursive.

5.2. Conversion d'image

Différents formats d'image par exemple JPG, GIF et TIF doivent être convertis au format pris en charge par le module de digestion. Comme les images sont supprimées après avoir été

digérées, le système peut conserver une copie compacte de chaque image (une vignette) nécessitant un espace de stockage minimal, afin d'être en mesure de présenter les résultats d'extraction à l'utilisateur. Le système devrait également permettre à l'utilisateur de voir l'image dans son état d'origine en les envoyant à son emplacement d'origine [5].

5.3. Validité d'image

Les images qui n'ont pas d'intérêt réel doivent être supprimées tels que les images publicitaires, de décoration, de navigation (flèches, retour à la maison, des cartes de site) et les logos d'entreprise pour ne pas être indexées [5].

5.4. Descripteurs d'image

Après la collection d'images et données à partir du Web, il est nécessaire de les classer en utilisant des caractéristiques (ou descripteurs) servant à indexer ces images et les récupérer.

Dans la recherche d'images basée sur du texte, les images sont annotées avec le texte entourant, ce dernier est analysé afin d'extraire un ensemble de mots-clés faisant référence à l'image. Les images peuvent être ensuite récupérées par mesure de correspondance entre les requêtes et les mots-clés de l'image.

Dans la recherche d'images basée sur le contenu, différentes caractéristiques de l'image sont utilisées : la couleur, forme, texture, l'emplacement.

Il est important de noter que les deux méthodes peuvent être utilisées en conjonction.

5.5. Similarité et correspondance

Pour la plupart des applications, les images sont dites conformes à la requête (image requête ou mots clés) si leurs descripteurs correspondent plus au moins à celles de la requête. Mais en matière de protection du droit d'auteur sur l'image, nous voulons trouver des copies de la même image, alors la correspondance peut être la plus appropriée, avec une certaine tolérance pour les imperfections [5].

5.6. Spécification de requête

Il existe 3 façons de faire une requête dans un système d'indexation et de recherche d'images :

- Des requêtes à base des mots clés.
- Des requêtes à base d'une esquisse.
- Des requêtes à base d'un exemple.

L'interface doit être aussi facile à utiliser que possible, tant au stade de la spécification de la requête et au stade de la présentation des résultats.

5.7. Récupération et raffinement

Le raffinement de la requête est assuré par le retour de pertinence. C'est le processus de réglage automatique de la requête à l'aide d'informations remontées par l'utilisateur sur la pertinence des images précédemment récupérées. La requête idéale est déplacée vers des exemples positifs et loin des exemples négatifs. Cela peut aider à identifier la requête qui est idéale à l'esprit de l'utilisateur. Des techniques comme la repondération des axes peuvent aider à déterminer les fonctions qui sont les plus importantes pour l'utilisateur [5].

5.8. Évaluation de la performance

En général, l'évaluation du système d'extraction des données est basée sur des essais sur une collection de documents, où l'utilisateur choisit un ensemble de requêtes avec un ensemble de scores de pertinence donné pour chaque paire requête-document, et un ensemble de paramètres d'évaluation tels que rappel² et précision³.

6. Présentation de LYCOS

Lycos est un moteur généraliste qui possède une interface permettant la recherche des images en plus bien sûr de la recherche de pages web. En outre, il permet la recherche des sons et des vidéos. Ces options peuvent ainsi être regroupées dans une catégorie globale de recherche multimédia [4].

Le catalogue et le moteur de recherche en ligne Lycos permettent d'effectuer des requêtes par mot clé et de les affiner à l'aide d'opérateurs booléens. Avec des requêtes plus fines, le moteur peut renvoyer plus de documents et ainsi augmenter le rappel [6].

6.1. La formulation syntaxique des requêtes

Les possibilités offertes pour l'expression d'une requête d'une manière générale sont diverses. En effet, il existe en total seize fonctionnalités parmi lesquelles on cite : les opérateurs syntaxiques que l'utilisateur peut employer (AND, OR, AND NOT, +, -, NEAR), la possibilité de faire une équation de recherche avec des parenthèses, la recherche d'une expression exacte, utilisation d'une requête avec troncature, gestion de la sensibilité à la

² Rappel : le ratio entre le nombre de documents pertinents retrouvés et le nombre total de documents pertinents dans la base

³ Précision : le nombre de documents pertinents retrouvés rapporté au nombre total de documents retrouvés

casse, le respect des accents, le filtrage du contenu, la restriction de sa recherche par certains critères (la langue, la date, le site, etc.) ou sur un champ particulier (le titre de la page, l'URL, etc.), et enfin la possibilité d'affiner sa recherche.

Lycos dispose sept fonctionnalités parmi les seize proposées précédemment. Il permet l'utilisation de l'opérateur AND, les opérateurs + et - ainsi que la recherche par une expression exacte. Les autres options proposées par ce moteur sont : le respect des accents, l'utilisation des parenthèses, et le filtre de contenu.

6.2.La fonction de filtrage

Vu l'importance que revêtent les filtres de contenu pour la recherche d'images, il semble intéressant de noter que tous les outils étudiés permettent d'appliquer aux résultats un filtre de contenu adulte. Il est désactivé par défaut sur *Lycos*, et peut être verrouillé le filtre par un mot de passe selon le choix d'utilisateur.

6.3.Présentation des résultats

La manière dont les résultats d'une recherche d'images sont présentés à l'utilisateur conditionne en grande partie le confort de celui-ci. L'information qui accompagne chaque image et qui concerne à la fois l'image elle-même (sa taille, son format, etc.) et sa source (de quel site est-elle issue ? que peut-on savoir sur ce site avant d'y accéder ? Etc.), ainsi que les modalités d'accès à la source de l'image. Lycos indique la taille des images (en pixels), leur poids (en kilo-octets), le nom de fichier et un lien direct vers l'image ou un lien vers la page web où elle apparaît.

6.4.La démarche suivie pour la recherche sur LYCOS

La démarche à suivre pour la recherche d'images par mots clés sur Lycos est la suivante :

- L'utilisateur lance sa requête sur le système.
- Le système envoie les mots clés comme requête aux moteurs de recherche et reçoit les URL des pages HTML trouvées.
- Recherche sur les pages HTML indiquées par les URL.
- Le système Analyse les pages HTML, extrait les URL des images incorporées, puis pour chaque image calcul un score en s'appuyant sur la concordance des mots clés utilisés par l'utilisateur avec les balises HTML de chaque image « TITLE », « SRC IMG », « ALT », « A HREF » (si l'image contient des références), etc.

- Le système sélectionne les images ayant obtenu un score supérieur à un certain seuil et les présente à l'utilisateur [7].

7. Conclusion

Dans ce chapitre nous avons présenté un aperçu général sur la recherche d'images sur le web. Une vision générale est donnée sur les types de recherche utilisés, les domaines d'applications, les enjeux et les techniques associées à la recherche des images sur le web. Puis, nous avons présenté le moteur de recherche LYCOS : la formulation syntaxique des requêtes, la fonction de filtrage et la présentation des résultats sur ce dernier.

Chapitre II : Recherche d'images par le contenu

- 1- Introduction
- 2- Historique de la recherche d'image par le contenu
- 3- Généralités de la recherche d'image
- 4- Les méthodes de recherche par le contenu
- 5- Conclusion

Chapitre II

Recherche d'images par le contenu

1. Introduction

Les pluparts de moteurs de recherche sur internet effectuent la recherche d'images en se basant sur des annotations textuelles ou sur le texte entourant les images. L'archivage des images est des séquences vidéos ne se fait qu'auprès d'une étape d'annotation textuelle à l'aide des mots clés [8]. Suite à aux inconvénients posés par cette dernière, il est nécessaire de recourir à d'autres méthodes, une de ces méthodes consiste à rechercher les images selon leur contenu.

La recherche d'images selon leur contenu permet de calculer une distance entre la requête image et chaque image dans la base de données en utilisant leurs caractéristiques de bas niveau (couleur, texture et forme) ou bien leurs contenu sémantique [9].

Dans ce chapitre nous allons présenter les différentes notions et concepts liés à la recherche d'images selon le contenu, ainsi que les différentes méthodes et algorithmes proposés dans ce contexte. Nous exposons avec plus de détails les deux méthodes de recherche : recherche par caractéristiques de bas niveau et par contenu sémantique.

2. Historique de la recherche d'image par le contenu

- ◆ 1980 : Annotation textuelle des images.
- ◆ 1990 : Indexation par la couleur, la forme et la texture.
- ◆ 1993 : Requête en utilisant la similarité des images.
- ◆ 1997 : Requête par images exemplaires et mesure de la pertinence (relevance feedback).
- ◆ 2000 : Apprentissage sémantique et adaptation à l'utilisateur.

- ✓ Enjeux actuels : Annotation interactive, formalisation de la connaissance (ontologies), grandes bases de données (problèmes de stockage, de vitesse d'accès, etc.), données hétérogènes, etc. [10].

3. Généralités de la recherche d'images

On assiste dans ces jours une croissance incontournable d'informations (texte, image et video) contenu dans le web dont voici quelques chiffres :

- ✓ 500.000 heures vidéo + 600.000 heures audio + 2.000.000 de photos.
- ✓ 82 Milliards de photos par année dans le monde.
- ✓ 390 Millions d'images indexées par Google.
- ✓ 4250 films commerciaux par année.

Cette croissance a conduit aux motivations suivantes :

- ✓ Conservation d'un patrimoine (culturel, scientifique, etc.).
- ✓ Valorisation en facilitant l'accès et l'exploration.
- ✓ Exploitation commerciale (photos numériques, TV numérique, etc.) [10].

4. Méthodes de recherche par le contenu

Les méthodes de recherche par le contenu permettent de rechercher des images d'une base d'images en fonction de leurs caractéristiques visuelles. Ces caractéristiques sont encore appelées caractéristiques de bas niveau (couleur, forme, texture, etc.). Aussi, il existe des méthodes de recherche par le contenu sémantique de l'image.

4.1. Les caractéristiques de bas niveau

Les images sont décrites par leurs caractéristiques visuelles (couleur, texture et forme) uniquement [11]. La *figure II.1* présente la méthode d'extraction de ses différentes caractéristiques.

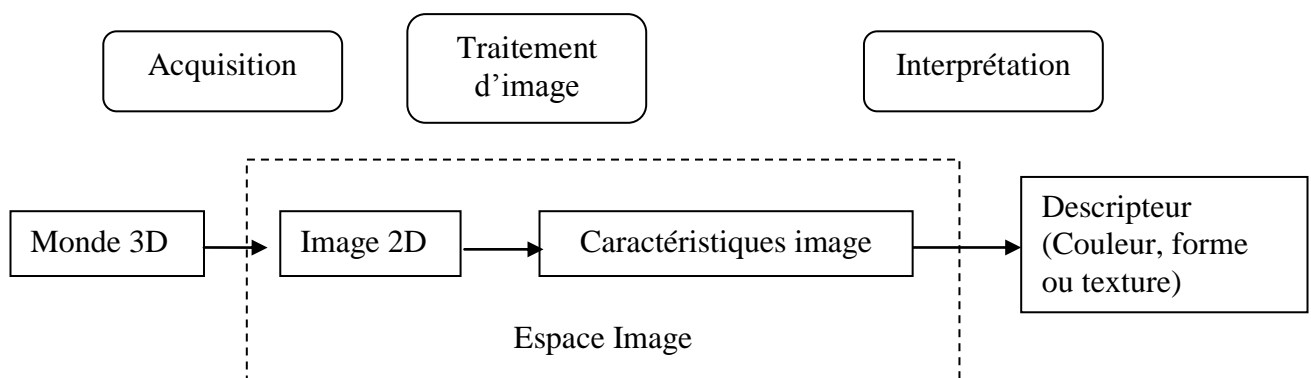


Figure II.1 : Extraction des descripteurs pour la recherche d'images.

4.1.1. La couleur

La couleur est très souvent le premier descripteur qui est employé pour la recherche d'images. Plusieurs travaux ont déjà prouvé qu'il s'agit d'un descripteur efficace. Il existe de

nombreuses possibilités d'attributs pour caractériser la couleur tels que l'histogramme, les moments couleur, etc. Et on trouve qu'avec la même méthode, si on change l'espace de couleur, il peut émerger des informations différentes de l'image.

1) L'histogramme

Une technique très utilisée pour la couleur est l'intersection d'histogrammes. Les histogrammes sont faciles et rapides à calculer, et robustes à la rotation et à la translation.

A) Les problèmes liés à l'utilisation d'histogramme

- Les histogrammes sont de grandes tailles, par conséquent, il est difficile de créer une indexation rapide et efficace.
- Ils ne possèdent pas d'informations spatiales sur les positions des couleurs.
- Ils sont sensibles à de petits changements de luminosité, ce qui est problématique pour comparer des images similaires, mais acquises dans des conditions différentes.
- ✓ Ils sont inutilisables pour la comparaison partielle des images (objet particulier dans une image), puisque calculés globalement sur toute l'image.

B) Les solutions des problèmes posés

Il existe deux solutions pour améliorer ces problèmes : l'une ajoute des informations spatiales à l'histogramme, l'autre utilise d'autres espaces de codage des couleurs qui se rapprochent plus de la perception d'humaine [8].

- ✓ La première approche ajoute des informations spatiales aux histogrammes. L'image est divisée en cinq blocs fixés superposés, puis, il faut extraire les trois premiers moments d'inertie de chaque bloc pour créer un vecteur de descripteurs. Ensuite, il faut ajouter la cohérence spatiale aux histogrammes. Un pixel est cohérent s'il appartient à une région validée par la segmentation et incohérent autrement (situé hors des régions segmentées). Les valeurs d'un histogramme sont divisées en deux classes en fonction des pixels correspondants à ces valeurs : classe cohérente et classe incohérente. La comparaison entre les deux histogrammes devient alors la comparaison entre les valeurs d'histogrammes dans les classes correspondantes.
- La deuxième approche recherche d'autres espaces de couleurs qui se basent sur la perception de couleur de l'humain.
 - ❖ L'espace RVB (Rouge Vert Bleu - en anglais RGB) est un espace de couleur utilisé couramment, par facilité, dans tous les systèmes de vision automatique,

mais il n'est pas forcément le mieux adapté. En effet, les trois composantes RVB sont très dépendantes les unes des autres. Un simple changement d'éclairage dans la scène modifie les trois composantes, alors que les objets de la scène n'ont pas changé de couleur, mais se sont simplement assombris ou éclairés.

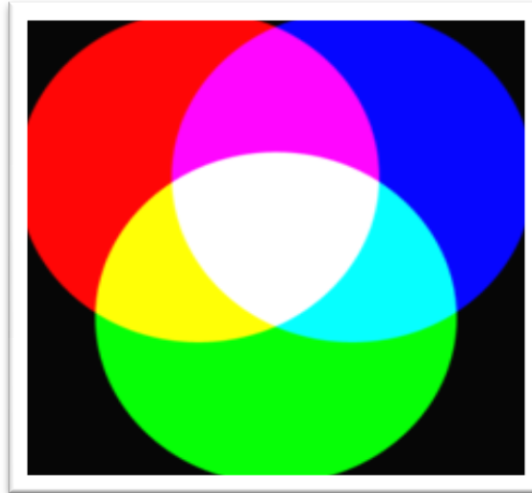


Figure II.2 : Présentation des couleurs dans l'espace RGB.

- ❖ L'espace TSV (Teinte Saturation Valeur - en anglais HSV) est le plus utile pour la segmentation et la reconnaissance et il a été prouvé un espace très fort dans le système de recherche des images. Parce qu'il est une représentation plus physique de la couleur [8]. Dans cet espace, il est possible pour un pixel de séparer : l'intensité du pixel (valeur) et la couleur du pixel (teinte et saturation). Cette séparation n'est pas disponible dans le codage RVB. En RVB les trois informations ne sont pas entièrement dé-corrélées.

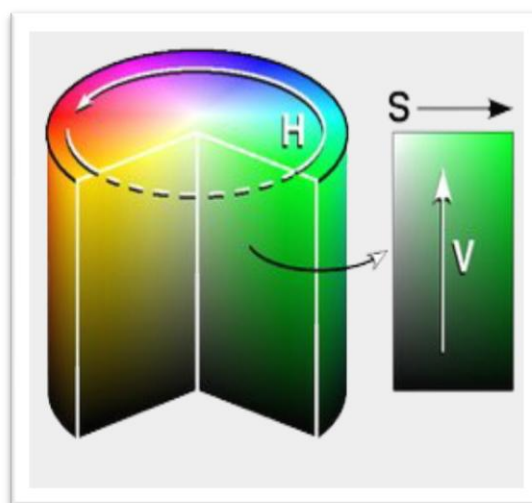


Figure II.3 : Présentation des couleurs dans l'espace TSV.

2) Les moments statistiques

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 dominantes caractéristiques de couleurs tels que l'espérance, la variance et d'autres moments.

L'espérance, la variance, les moments d'ordre 3 peuvent également être calculés sur chaque composante couleur par la formule suivante :

$$E_i = \frac{1}{N} \sum_{j=1}^N P_{ij} \quad (\text{II.1})$$

$$\delta_i = \left(\frac{1}{N} \sum_{j=1}^N (P_{ij} - E_i)^2 \right)^{1/2} \quad (\text{II.2})$$

$$S_i = \left(\frac{1}{N} \sum_{j=1}^N (P_{ij} - E_i)^3 \right)^{1/3} \quad (\text{II.3})$$

Où E_i est l'espérance, δ_i est la variance, S_i est le troisième moment qui correspond à la composante couleur i (i est R, V, B de l'espace RVB ou T, S, V de l'espace TSV) [8].

4.1.2. La texture

La deuxième primitive visuelle est la texture de l'image. La méthode la plus connue pour analyser la texture est la matrice de co-occurrences de Haralick.

Il existe aussi d'autres méthodes pour analyser les textures dont celles basées sur : la transformée en ondelettes de Gabor, les caractéristiques de Tamura, Word de composition, Fractal, etc.

Après avoir appliqué la transformation de Gabor sur une image, une région de texture est caractérisée par la moyenne et la variance des coefficients de transformation. Un vecteur de caractéristiques est construit en utilisant ces caractéristiques comme composants. Les caractères symboliques calculés à partir de la transformée en ondelettes de Gabor sont notamment très utilisés à l'heure actuelle pour caractériser les textures. L'intérêt principal des coefficients de Gabor est leur invariance aux rotations et mise à l'échelle [8]. La figure II.4 présente la recherche d'image par contenu en se basant sur la texture.

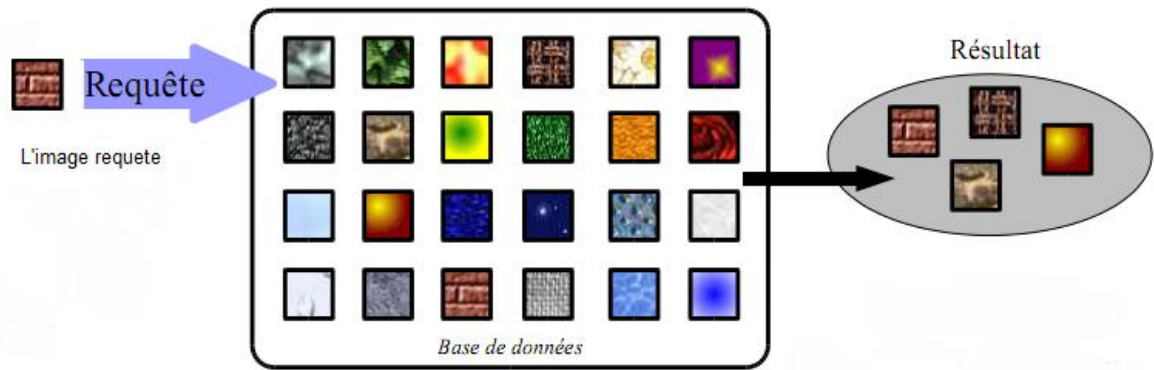


Figure II.4 : Recherche d'image en se basant sur la texture de l'image.

4.1.3. La forme

La forme est une caractéristique visuelle importante et elle est l'une des caractéristiques de base pour décrire le contenu d'une image. Cependant, la description des formes est une tâche difficile, surtout dans un contexte de « segmentation faible » et dans un contexte où on recherche la similitude et non pas l'exactitude. Ainsi, dans les systèmes de recherche récents, peu entre elles emploient la forme comme critère de recherche.

Les caractéristiques de forme sont extraites à partir des régions dans les images qui ont été segmentées d'une manière quelconque [9].

1) Les différentes approches de calcul des caractéristiques locales

Deux approches différentes peuvent être employées pour calculer les caractéristiques locales :

- a) La première approche consiste à diviser une image en utilisant une grille, ensuite les caractéristiques de chaque bloc sont calculées pour chaque case de cette grille. Par exemple dans la *figure II.5*, l'image est divisée en 12 zones de même taille et les caractéristiques sont calculées pour chacune de ces zones.



Figure II.5 : Découpage en grille de l'image.

b) La deuxième approche consiste à segmenter l'image pour la diviser en zones locales plus proches des objets constituant l'image et ensuite calculer les caractéristiques pour chacune des régions extraites. La segmentation est une étape délicate, à cause du problème de choix d'une méthode de segmentation qui convient pour toutes les images de la base d'images. Cependant, la division d'une image en régions ou objets d'intérêt est souvent nécessaire pour pouvoir s'attaquer ensuite à l'extraction d'informations sémantiques à partir de l'image. Plutôt que de viser une segmentation exacte (en termes d'interprétation de l'image), il est préférable de procéder par « groupement de pixels ». Dans ce cas, il s'agit simplement d'oublier le découpage précis en termes d'objets, mais plutôt de découper l'image en régions similaires du point de vue caractéristiques mesurées sur l'image (voir *figure II.6*).



Figure II.6 : Découpage par la segmentation division-fusion de l'image.

4.2. Les méthodes sémantiques

Les systèmes basés sur les caractéristiques de bas niveau ont des limitations parce que ces systèmes comparent seulement la similarité des caractéristiques de l'image. Dans les systèmes de recherche en sémantique, l'analyse de l'image est effectuée en termes d'objets et de contenu, et non pas seulement en termes de statistiques sur les couleurs, les textures ou autres caractéristiques bas niveau. La sémantique elle-même n'est pas inscrite dans l'image, mais se trouve ailleurs. Il faut donc rechercher les sources extérieures permettant l'accès aux clés de décodage sémantique de l'image. Dans une image, la sémantique dépend de deux éléments:

- ✓ Du niveau de connaissances et de la perception que possède l'observateur.
- ✓ De l'objectif poursuivi par l'utilisateur de cette image lorsqu'il la regarde.

Donc, on a deux approches pour retrouver cette sémantique :

- ✓ Rechercher des moyens pour connecter (ou lier) la connaissance sémantique humaine et l'apparence de l'image (les caractéristiques extraites de celle-ci).
- ✓ Rechercher des méthodes pour comprendre l'objectif de l'utilisateur, le sens de sa Requête [9].

4.2.1. Les travaux réalisés

1) Le modèle basé sur l'intelligence artificielle et la reconnaissance des formes

A) Description du modèle

Le principal objectif de ce modèle est de développer une méthodologie pour la représentation et l'extraction des connaissances à partir des images en utilisant les techniques de l'intelligence artificielle, afin de ramener la recherche d'image à une opération d'inférence utilisant les connaissances récoltées dans les images. La solution consistait donc à décrire d'une façon automatique le contenu des images. Ceci permettrait de dépasser le fossé sémantique, en d'autres termes le problème lié à la subjectivité des utilisateurs de la base.

B) Architecture du modèle

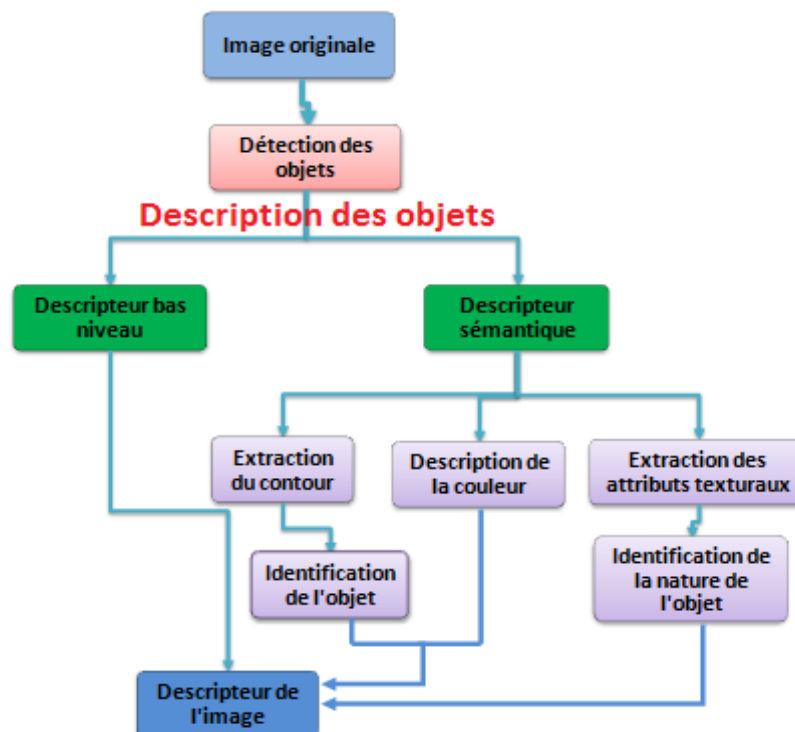


Figure II.7 : Architecture du modèle basé sur l'intelligence artificielle et la reconnaissance des formes.

Comme la montre la **figure II.7**, le modèle utilise deux types de descripteur pour l'indexation des images :

- ◆ Un descripteur sémantique qui permet la recherche d'images par la sémantique.
- ◆ Un descripteur de bas niveau qui servira comme fonction d'appariement pour ordonner le résultat de la recherche. Le descripteur de bas niveau peut aussi être utilisé pour la recherche en termes de similarité visuelle ou lorsque l'utilisateur juge que le résultat de la recherche sémantique n'est pas satisfaisant.

Les étapes principales de ce modèle sont les suivantes :

- 1) La première étape consiste à séparer les objets de l'image : pour cela, il est nécessaire d'utiliser un algorithme de segmentation d'image.
- 2) La description de bas niveau de chaque région de l'image : Ce descripteur servira à établir un ordre de similarité pour le résultat de la recherche ou encore à faire une recherche d'image en se basant sur la similarité visuelle.
- 3) La description sémantique de chaque objet de l'image qui consiste à :
 - a) Identification de l'objet : il faut dans un premier lieu détecter le contour de l'objet, puis utiliser un algorithme d'apprentissage pour la reconnaissance des formes. Ce descripteur a pour rôle de générer des labels comme : table, avion, personne, etc.
 - b) Nommage de la couleur de l'objet. Ex. vert, noir, rouge, etc.
 - c) Identification de la nature de l'objet : on commence par extraire les attributs texturaux de l'objet, puis on utilise un algorithme d'apprentissage pour la reconnaissance du modèle. Ce descripteur a pour rôle de générer des descriptions comme : objet en bois, en marbre, en brique, etc.
- 4) Enfin, construire un vecteur descripteur de l'image qui va regrouper la description de bas niveau et la description sémantique de l'image [12].

C) Résultats

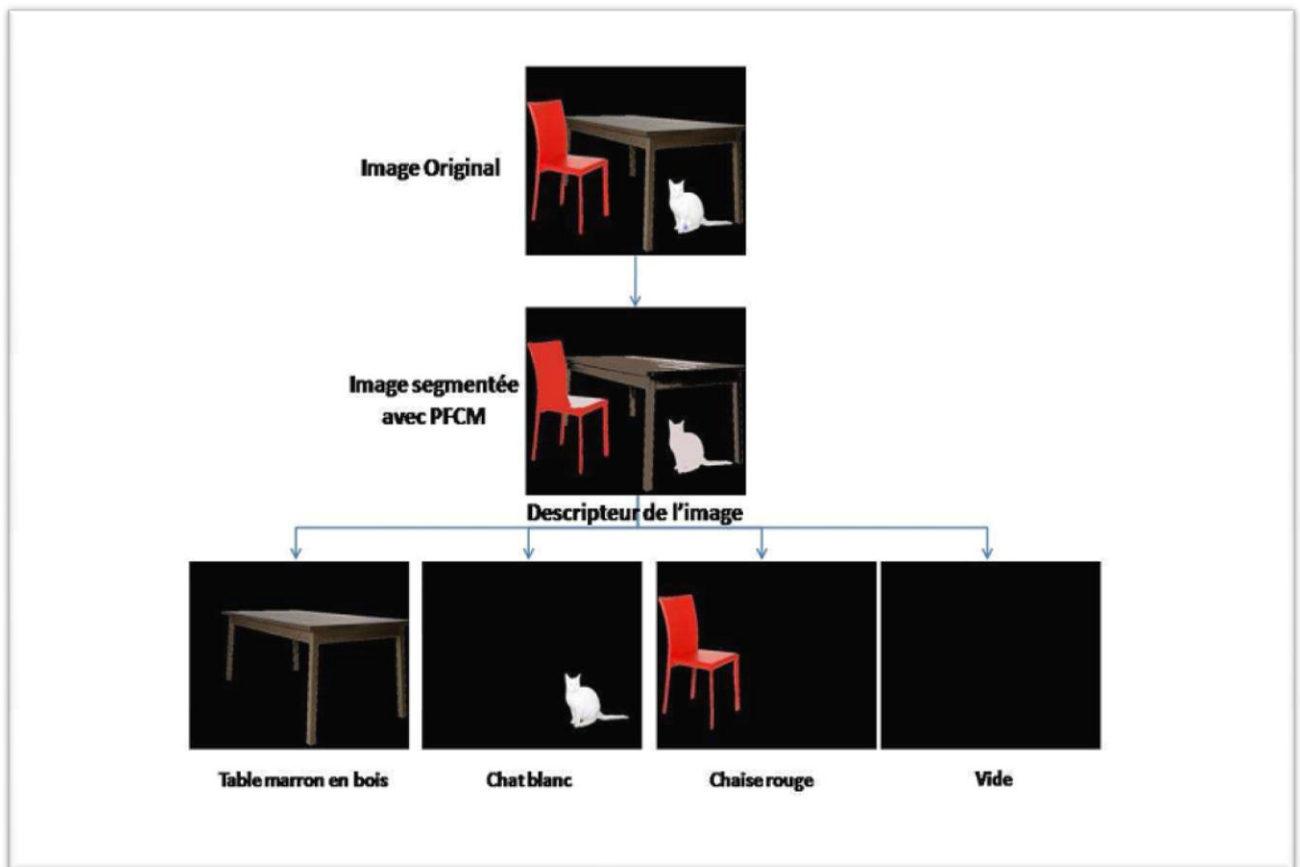


Figure II.8 : Résultat du modèle basé sur l'intelligence artificielle et la reconnaissance des formes.

2) Retour à la pertinence

Une voie de plus en plus utilisée pour pallier le manque de connaissances dans un système de vision par ordinateur est de concevoir des systèmes semi-automatiques. Cette voie existant déjà pour d'autres domaines, mais devient encore plus évidente dans le domaine de l'indexation et recherche d'images. Ici, on suppose que l'humain maîtrise la sémantique et qu'avec l'aide de quelques interactions simples et ciblées (par exemple, l'identification d'exemples positifs et négatifs à partir d'un premier retour du programme), le système peut prendre cette sémantique. Il est connu qu'une même image requête présentée par l'utilisateur peut avoir différentes significations selon les besoins immédiats de cet utilisateur (objet avant-plan, arrière-plan). La théorie prédominante est celle du retour de pertinence (relevance feedback en anglais) déjà utilisée dans le domaine de la recherche d'informations. Dans ce cas-ci, la sémantique n'est pas dans le système informatique, mais elle est extraite implicitement à chaque coup via l'interaction et l'analyse des retours.

Le processus de retour de pertinence est une technique de recherche d'images permettant, par la communication avec l'utilisateur, d'extraire de façon rapide et conviviale son objectif. Dans ce contexte, une interface graphique 2D adaptée au problème de la recherche d'images permettant une communication bidirectionnelle : du système vers l'utilisateur pour lui visualiser les résultats courants de la recherche et de l'utilisateur vers le système pour qu'il fournisse des informations de type retour de pertinence pour raffiner sa requête. Le retour de pertinence est uniquement sous forme de validation d'images résultats (ou exemples positifs) [13] (voir *figure II.9*).

La *figure II.10* donne un exemple de retour de pertinence dans le système d'indexation et recherche d'images. Dans cet exemple, le processus de retour de pertinence est répété deux fois ce qui a permis d'augmenter le nombre d'images pertinentes [8].

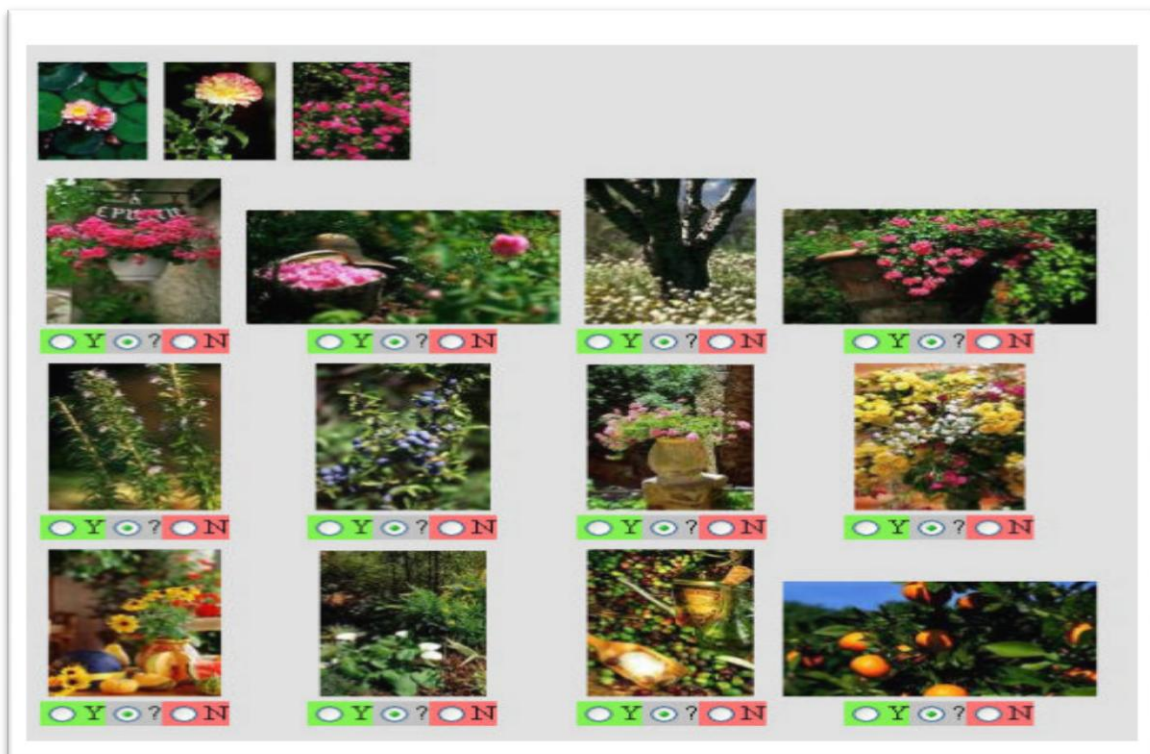


Figure II.9 : Exemple du processus de retour de pertinence.

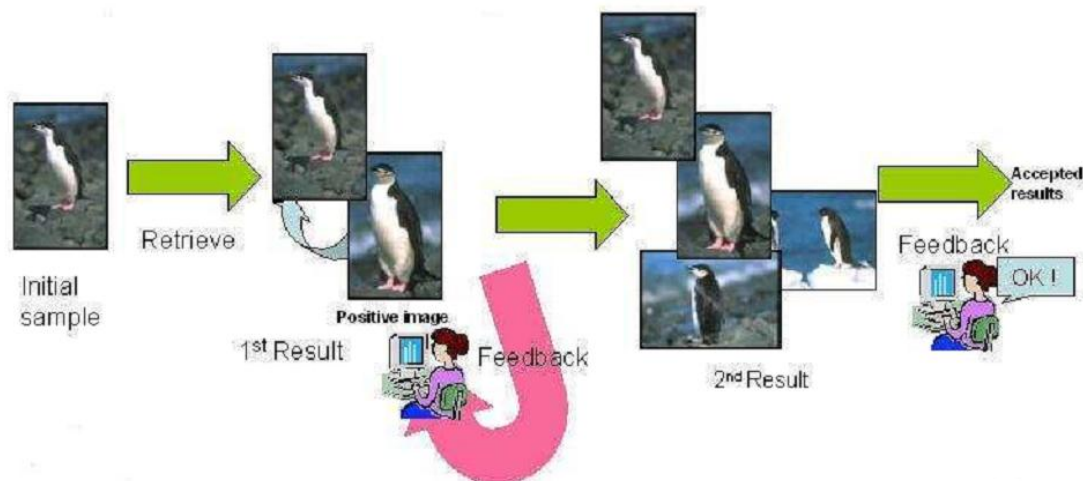


Figure II.10 : Un exemple de retour de pertinence dans le CBIR.

3) Ontologie (définition des concepts)

La définition des concepts est la définition de termes sémantiques utilisée pour décrire l'image, et la définition en termes de caractéristiques extraites de l'image [9].

Il existe plusieurs mécanismes de représentation de connaissances issus du domaine de l'intelligence artificielle. Par exemple, le problème consiste à rechercher des objets inclus dans des images comme une : voiture, un tigre, etc. L'homme peut faire cela facilement. Pour l'intelligence artificielle, la question est la suivante : Comment prendre les connaissances d'un domaine et les présenter sous une forme compréhensible par l'ordinateur lui permettant de reconnaître des objets ? Comment faut-il faire pour relier la connaissance haute niveau et l'apparence de l'image ?

Pour présenter des connaissances, les ontologies peuvent être utilisées dans le domaine de présentation des concepts.

Depuis quelques années, une direction de recherche vise à combiner des techniques conventionnelles avec une ontologie. Dans cette direction, on utilise des techniques d'apprentissage comme l'arbre décision, réseaux de neurones, etc. pour connecter des régions (objets) correspondant aux concepts définis dans l'ontologie. Puis, on peut utiliser des relations entre les concepts pour vérifier le résultat. Où, à partir de la requête d'utilisateurs, en utilisant l'ontologie, la requête peut être précisée, c'est-à-dire trouver des concepts plus détaillés. Puis, la recherche se fait à base des caractéristiques liées aux concepts détaillés.

▪ Comment définir les concepts et reconnecter les caractéristiques aux des concepts

Une ontologie peut être définie comme des spécifications d'un vocabulaire de représentation pour un domaine partagé du discours qui peut inclure des définitions de classes, des relations, de fonctions et d'autres objets.

La connaissance dans les ontologies est principalement formalisée en utilisant les cinq types de connaissance : classe, relations, fonctions, axiomes et instances. L'ontologie peut être construite comme suit [9] :

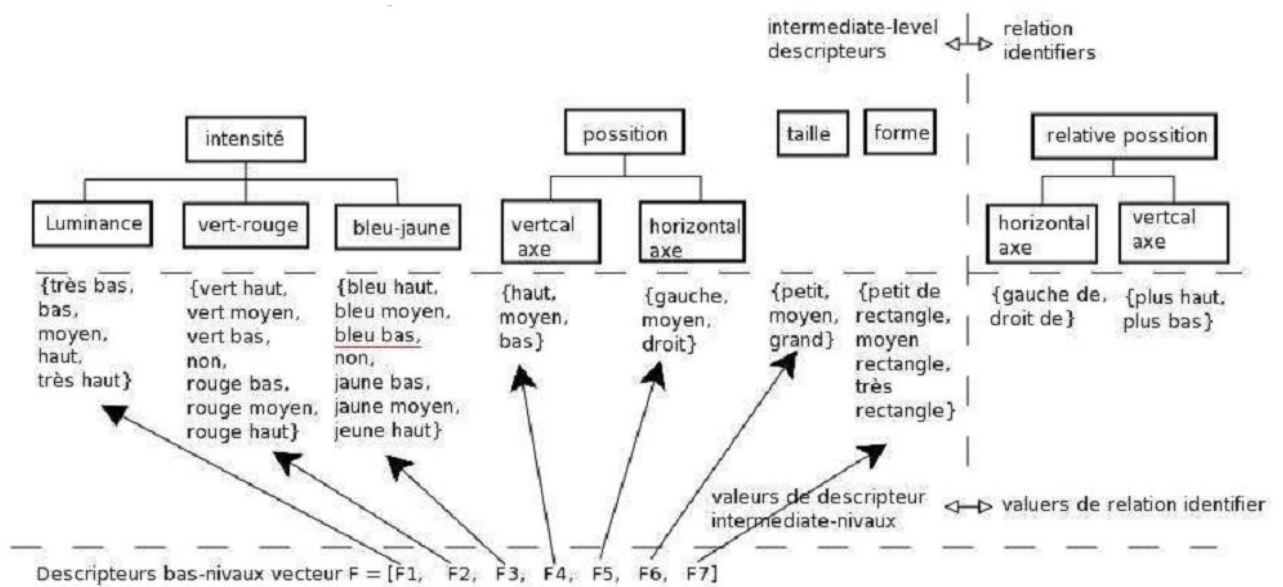


Figure II.11 : Exemple de construction d'une ontologie.

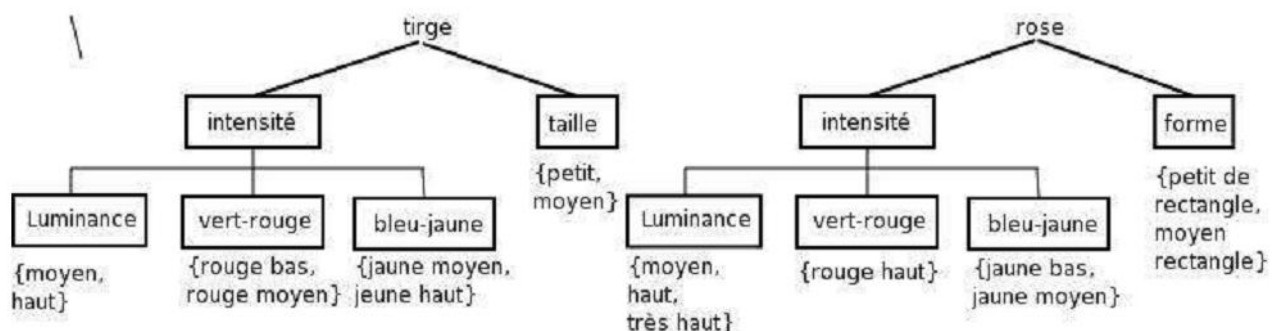


Figure II.12 : Cas d'utilisation d'une ontologie.

Les descripteurs bas niveau sont la couleur, la position, la taille et les formes des régions.

Les descripteurs de niveaux intermédiaires forment un simple vocabulaire « objet ontologie » associé aux descripteurs de bas niveau. Ces objets ontologies sont utilisés pour permettre une

définition qualitative de concepts hauts niveaux que l'utilisateur requiert dans le concept humain.

- *Voici un exemple d'une ontologie « Annotation d'images biomédicales et raisonnement »* : qui facilite le diagnostic par exploitation d'une ontologie du domaine du cancer du sein [14] (voir *Figure II.13*).

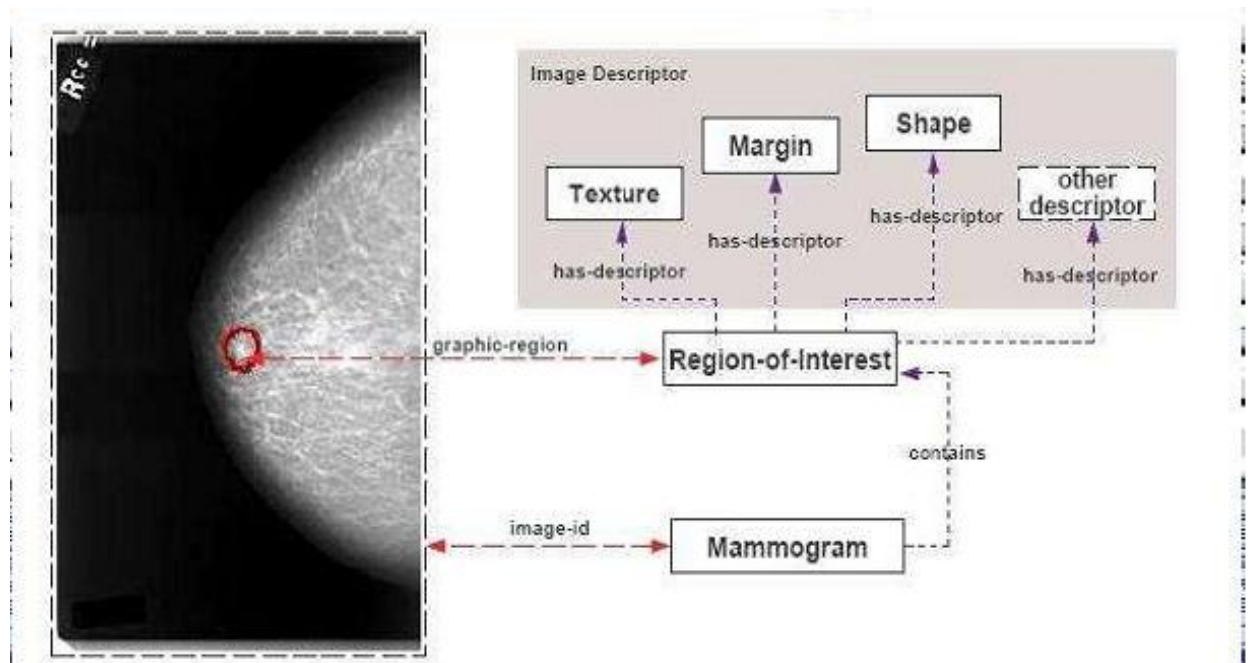


Figure II.13 : Exemple d'une ontologie « Annotation d'images biomédicales et raisonnement ».

5. Conclusion

Dans ce chapitre, nous avons présenté un bref historique sur la recherche d'images par le contenu, puis les méthodes de recherche les plus utilisées, ainsi que les principales caractéristiques de bas niveau sur lesquelles se basent ces méthodes. Enfin, nous terminons cette partie par la présentation des méthodes sémantiques soutenues par quelques exemples illustrant les travaux réalisés dans ce contexte (le modèle basé sur l'intelligence artificielle et reconnaissance des formes, le retour de pertinence et les ontologies).

Chapitre III : Reconnaissance des formes et détection d'objets

A. Introduction

Partie I: La reconnaissance des formes

I.1- Introduction

I.2- Le rôle de reconnaissance des formes

I.3- Schéma général d'un système de Reconnaissance de Formes

I.4- Méthode RDF

I.5- Champs d'Applications

Partie II: Détection d'objets

II.1- Définition

II.2- Les algorithmes de reconnaissance d'objet

II.3- Domaines d'application

II.4- Exigence et contraintes

II.5- Classification des méthodes de reconnaissance

B- Conclusion

Chapitre III

Reconnaissance des formes et détection d'objet

1. Introduction

La reconnaissance des formes (RdF) a pour but de concevoir de systèmes hard/soft capables de percevoir et d'interpréter des informations captées dans le monde réel. La conception des systèmes automatiques ou semi-automatiques à l'image de l'homme qui reconnaissent les formes qu'on leur présente est loin d'être résolue. L'homme est un système biologique le plus parfait des systèmes de RdF. La reproduction des facultés humaines, des fonctions physiques typiquement humaines sur machines nécessite une bonne représentation des modules de perception, d'analyse, de représentation et d'interprétation. Quant à la compréhension automatique des informations, elle relève de l'intelligence artificielle [15].

Dans ce chapitre, nous allons présenter deux parties, l'une se consacre à la reconnaissance des formes et l'autre à la détection d'objets. Dans la première partie nous définissons la RdF, ses différents rôles, son schéma général, ses approches, ses méthodes, et finalement nous citons les différents champs d'application. Dans la deuxième partie nous allons présenter les différents algorithmes de reconnaissance d'objet, les domaines d'applications, exigences et contraintes, et finalement nous terminons par une classification des méthodes de reconnaissance.

Partie I : La reconnaissance des formes

I.1. Définition des RDF

La reconnaissance des formes est une discipline qui a pris naissance au milieu des années cinquante. Elle se considère comme une science d'élaboration d'algorithmes permettant de classer des objets dont l'aspect varie par rapport à un objet type. Son objectif est la conception et l'analyse des algorithmes de classification applicables aux présentations d'images, de sons ou de signaux de toute nature. Les disciplines les plus proches de la reconnaissance des formes sont :

- **Pour les mathématiques** : Le calcul des probabilités, la logique, l'analyse combinatoire, la géométrie intégrale, le traitement du signal.
- **Pour l'informatique** : La grammaire formelle, la théorie de la démonstration, les systèmes à base de connaissance.

Une des tâches importantes de la reconnaissance des formes est de caractériser la forme, c'est-à-dire de lui trouver une description qui la singularise, mais aussi la rapproche de ses formes voisines et qui l'éloignent de ses fausses semblables [16].

I.2. Le rôle de reconnaissance des formes

- La reconnaissance des formes vise la perception et l'interprétation de l'environnement.
- Une phase préliminaire consiste à numériser des éléments du monde réel (acquisition des données).
- Étant donnée une information riche et non structurée (par exemple une image), la RdF a pour but de lui associer une signification symbolique (par exemple une étiquette).
- La RdF peut être perçue comme un procédé méthodique de réduction d'information, pour ne conserver que l'interprétation pertinente [17].

I.3. Schéma général d'un système de reconnaissance de formes

Un problème de RdF est généralement décrit comme une chaîne de transformation mettant en relation trois espaces différents [15]:

- L'espace de mesure.
- L'espace de représentation.
- L'espace de décision ou d'interprétation.

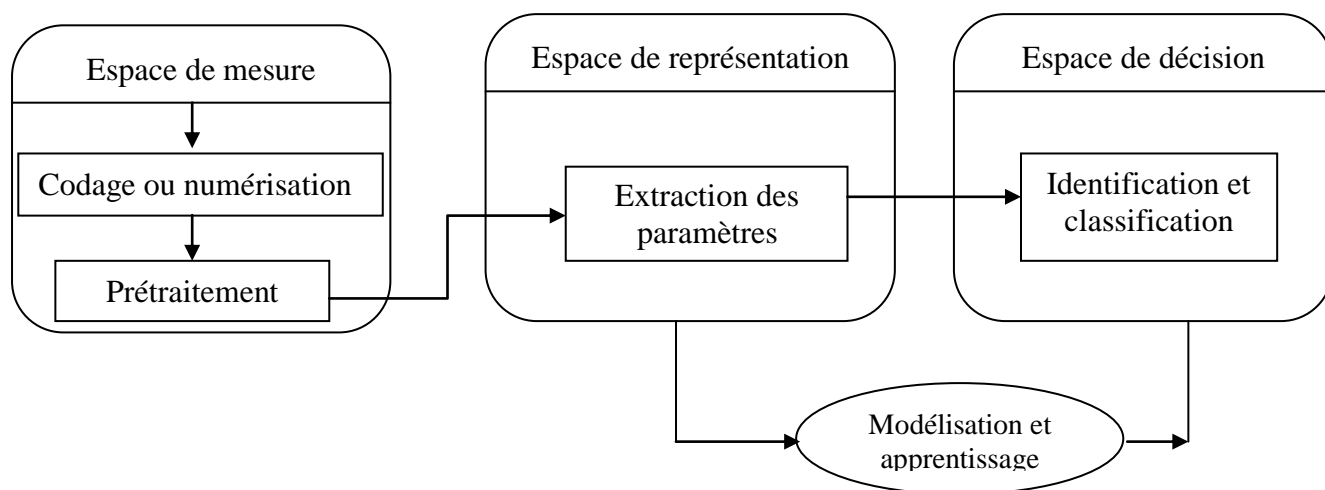


Figure III.1 : Schéma général d'un système de reconnaissance de forme.

I.3.1. L'espace de mesure

L'espace de mesure est une représentation du monde physique (réel), qui est un espace analogique de dimensions infinies. Les objets y sont décrits de différentes façons avec une multitude de propriétés difficiles à prendre en compte dans leur ensemble lors d'une reconnaissance. Il comprend deux étapes :

- **L'étape de codage ou numérisation** : elle consiste à transformer, par l'association d'un capteur, un ensemble de données analogiques en données numériques manipulables par un ordinateur. Cette transformation doit se faire de la manière la plus fidèle possible en conservant les propriétés essentielles de l'objet physique.
- **L'étape de prétraitement** : les données brutes sont souvent bruitées, elles contiennent des informations parasites et n'explicitent pas les informations utiles pour la reconnaissance. Le rôle de cette étape est de préparer ces données à la phase de l'extraction de paramètres en éliminant le bruit (souvent dû aux conditions d'acquisition) en corrigeant les erreurs, en normalisant et réduisant les données.

I.3.2. L'espace de représentation

C'est la représentation par une description adéquate des informations provenant de l'espace de mesure, il correspond à l'étape d'extraction des paramètres. Lors de cette étape, il y a un calcul ou une détermination des propriétés caractéristiques de la forme notée : $X_1, \dots, X_i, \dots, X_p$ qui prennent leurs valeurs respectivement dans des domaines de variation : $R_1, \dots, R_i, \dots, R_p$. L'espace des représentations est noté R qui représente le produit cartésien $R_1 \times \dots \times R_i \times \dots \times R_p$.

Soit Π une population d'individus, d'objets ou d'exemples concernés par le problème de la RdF, la fonction qui permet d'associer à chaque exemple sa représentation noté χ est définie comme suit :

$$\chi : \Pi \rightarrow R = R_1 \times \dots \times R_i \times \dots \times R_p \quad (\text{III.1})$$

$$\pi \rightarrow \chi(\pi) = (X_1(\pi), \dots, X_i(\pi), \dots, X_p(\pi))$$

Où : $X_i(\pi)$ représente la valeur prise par l'attribut X_i pour l'exemple π .

I.3.3. L'espace de décision ou d'interprétation

L'interprétation d'une représentation est l'étiquette, la classe, ou le concept qui lui est associé. On parle généralement de classe. Les classes sont représentées par une variable notée Y qui

prend ses valeurs dans un ensemble C discret et de cardinal fini C . C est appelé l'espace d'interprétation.

$$\begin{aligned} Y: \Pi &\rightarrow C = \{Y_1, Y_2, Y_c\} \\ \pi &\rightarrow Y(\pi) \end{aligned} \quad (\text{III.2})$$

L'objectif de la RdF est de construire une procédure de prévision ϕ qui permette d'expliquer et de prédire la classe de tous les exemples de la population Π . Cette procédure, encore appelée identification, opérateur de RdF ou fonction de classement, est une application de l'espace de représentation dans l'espace d'interprétation :

$$\begin{aligned} \varphi: R &\rightarrow C \\ \chi(\pi) &\rightarrow Y_i \end{aligned} \quad (\text{III.3})$$

I.4. Méthode RDF

La reconnaissance de motifs peut être effectuée au moyen de divers algorithmes d'apprentissage automatique tels:

- Un réseau de neurones.
- Une analyse statistique.
- L'utilisation de modèles de Markov cachés.
- Une recherche d'isomorphisme de graphes ou sous-graphes.

Les formes recherchées peuvent être des formes géométriques, descriptibles par une formule mathématique, telles que :

- Cercle ou ellipse.
- Droite.

Elles peuvent aussi être de nature plus complexe:

- Lettre.
- Chiffre.
- Empreinte digitale.

Les algorithmes de reconnaissance peuvent travailler sur des images en noir et blanc, avec les contours en blanc des objets se trouvant dans l'image. Ces images sont le fruit d'algorithmes de détection de contours. Ils peuvent aussi travailler sur des zones de l'image prédéfinies issues de la segmentation de l'image [18].

1) Méthode Bayésienne

L'inférence bayésienne est la démarche logique permettant de calculer ou réviser la probabilité d'une hypothèse. Cette démarche est régie par l'utilisation de règles strictes de combinaison des probabilités, desquelles dérive le théorème de Bayes. Dans la perspective bayésienne, une probabilité n'est pas interprétée comme le passage à la limite d'une fréquence, mais plutôt comme la traduction numérique d'un état de connaissance (le degré de confiance accordé à une hypothèse, par exemple le théorème de Cox-Jaynes) [Web 1].

2) Classifieur linéaire

En apprentissage automatique, le terme de classifieur linéaire représente une famille d'algorithmes de classement statistique. Le rôle d'un classifieur est de classer dans des groupes (des classes) les échantillons qui ont des propriétés similaires, mesurées sur des observations. Un classifieur linéaire est un type particulier de classifieur, qui calcule la décision par combinaison linéaire des échantillons [Web 2].

3) Réseau de neurones

Un réseau de neurones artificiels est un modèle de calcul dont la conception est très schématiquement inspirée du fonctionnement des neurones biologiques. Les réseaux de neurones sont généralement optimisés par des méthodes d'apprentissage de type probabiliste, en particulier bayésiens. Ils sont placés d'une part dans la famille des applications statistiques, qu'ils enrichissent avec un ensemble de paradigmes permettant de générer des classifications rapides (réseaux de Kohonen en particulier), et d'autre part dans la famille des méthodes de l'intelligence artificielle auxquelles ils fournissent un mécanisme perceptif indépendant des idées propres de l'implémenter, et fournissant des informations d'entrée au raisonnement logique formel [Web 3].

4) SVM (Support Vector Machine)

Les séparateurs à vastes marges sont des classifieurs qui reposent sur deux idées clés, qui permettent de traiter des problèmes de discrimination non linéaire, et de reformuler le problème de classement comme un problème d'optimisation quadratique.

La première idée clé est la notion de *marge maximale*. La marge est la distance entre la frontière de séparation et les échantillons les plus proches. Ces derniers sont appelés *vecteurs supports*.

Afin de pouvoir traiter des cas où les données ne sont pas linéairement séparables, la deuxième idée clé des SVM est de transformer l'espace de représentation des données d'entrées en un espace de plus grande dimension (possiblement de dimension infinie), dans lequel il est probable qu'il existe un séparateur linéaire [Web 4].

I.5. Champs d'Applications

Les applications attachées classiquement à cette discipline sont les suivantes :

- 1) **Reconnaissance de la Parole** : liée à plusieurs disciplines comme le décodage acoustico-phonétique, la reconnaissance de mots, la reconnaissance de phrases, la reconnaissance du locuteur et compréhension du dialogue oral homme-machine. Parmi les applications les plus courantes, on peut citer la commande vocale, la dictée automatique, la traduction en temps réel de langues étrangères et la rééducation de malentendants.
- 2) **Reconnaissance de l'écriture** : La reconnaissance de l'écriture est attachée à la reconnaissance des caractères manuscrits et imprimés, la reconnaissance de mots et de phrases, la reconnaissance du scripteur et la reconnaissance de documents. Parmi les applications de l'écriture on a le tri automatique du courrier par lecture et reconnaissance des adresses, l'authentification des chèques bancaires, la saisie et l'archivage des documents.
- 3) **La Vision** : La vision comprend le traitement, l'analyse et l'interprétation des images. Les applications dans ce domaine sont assez nombreuses. On peut citer la sûreté avec la reconnaissance des empreintes digitales, la médecine avec l'analyse d'images radiographies ou d'échographies, l'industrie avec l'analyse des défauts dans les pièces de fabrication où l'identification d'objets, la surveillance de processus en robotique, reconnaissance des visages, la géophysique avec l'analyse des images satellites pour les prévisions météorologiques, l'armée avec la surveillance, le guidage de cible, etc. Ces applications nécessitent une adaptation importante des unités de calcul et une spécialisation du matériel d'acquisition des formes à traiter [19].

Partie II : Détection des objets

II.1. Définition

La reconnaissance d'objet est l'un des principaux domaines de traitement d'images et de reconnaissance des formes. L'expression détection ou reconnaissance d'objet est largement utilisée, souvent elle désigne une exploration d'une ou plusieurs images est effectuée pour

déterminer quels objets sont présents dans l'image et où sont-ils ? Dans ce contexte, quelques connaissances sur les objets recherchés sont nécessaires (un modèle qui a été créé d'avance). Donc, à base de cette connaissance, l'image est examinée pour rechercher le(s) instance(s) de l'objet et localiser sa (ses) position(s). Chaque application de reconnaissance ou détection d'objet dispose de ses spécificités, exigences et contraintes sous lesquelles elle fonctionne. Pour répondre à toutes ces exigences et contraintes, beaucoup d'algorithmes ont été proposés au fil du temps [20].

II.2. Les Algorithmes de reconnaissance d'objets

Les algorithmes de reconnaissance d'objets sont classés en différentes catégories :

- 1) **Méthodes généralistes** : Les méthodes généralistes font l'apprentissage à partir d'une base d'apprentissage.
- 2) **Méthodes de recherche basées sur la transformation** : Ces méthodes cherchent à prouver l'apparition d'un modèle spécifique dans une position spécifique par exploration de l'espace des transformations possibles entre le modèle et les données de l'image.
- 3) **Méthodes guidées par des modèles géométriques** : Ces méthodes sont dérivées des méthodes de séparation de données CAD¹ et visent à diviser l'objet (le modèle) en parties ayant des relations géométriques spécifiques les uns avec les autres, la reconnaissance est établie en faisant la correspondance entre les différentes parties de l'image et le modèle de l'objet.
- 4) **Méthodes basées sur un modèle 3D** : Ces méthodes visent à localiser la position d'un objet 3D dans une image 2D, essentiellement par la recherche de caractéristiques qui sont invariables selon l'angle de vue.
- 5) **Méthodes de construction de contour flexible et contour active** : Ces méthodes visent à chercher dans l'image une courbe (contour) comparable à celle de l'objet à reconnaître.
- 6) **Méthodes basées sur des descripteurs** : Le modèle est représenté comme une collection de descripteurs. La reconnaissance s'effectue par l'exploration de l'image à la recherche d'une ou plusieurs instances correspondantes à cette collection.

II.3. Domaines d'application

La seule manière de montrer la diversité des cas d'utilisation d'application de reconnaissance d'objets est de citer leurs domaines d'application, ces domaines incluent le domaine industriel

¹ Computer Aided Design

(dans ce cas c'est le terme computer vision qui est souvent utilisé), militaire, médicale, les applications de sécurité, de contrôle, etc.

- 1) **Localisation de la position d'objet (Position measurement) :** Généralement dans les applications industrielles, il est indispensable de localiser avec précision la position des objets. Cette information (la position) est nécessaire pour prendre, traiter, déplacer ou placer des objets dans un environnement de production, Ex : il est nécessaire de reconnaître et localiser avec précision la position des composants électroniques pour les souder sur le circuit imprimé dans la production des téléphones portable micro-ordinateurs, etc.
- 2) **Inspection :** L'utilisation des systèmes de vision pour le contrôle de qualité dans un environnement de production est une application classique de la vision par ordinateur. Typiquement la surface des objets produits est inspectée pour détecter ceux défectueux.
- 3) **Trie (Sorting) :** Dans ce type d'applications de reconnaissance d'objet l'objectif est de trier les objets selon des caractéristiques bien définies, par exemple : Tri des lettres selon leur taille.
- 4) **Comptage :** Quelques cas d'utilisation d'applications de reconnaissance d'objets exigent la détermination du nombre d'occurrence d'un objet dans une image. Par exemple : Détermination du nombre d'un organisme biologique dans une image microscopique.
- 5) **Détection d'objet :** Dans ce cas d'utilisation, les données de l'image sont comparées avec ceux du modèle représentant les objets recherchés pour identifier l'objet ou les objets figurants sur l'image. Dans ce contexte, l'objectif est donc de confirmer l'existence d'objets sur l'image.
- 6) **Catégorisation de scène :** L'objectif de la catégorisation n'est pas de reconnaître des objets dans l'image, mais consiste à classer cette dernière selon une catégorie (l'objectif n'est pas de reconnaître le contenu de l'image mais de reconnaître le sujet de la scène).
- 7) **Recherche d'images (Image Retrieval) :** Ces applications de reconnaissance sont utilisées pour rechercher des images semblables à une image requête selon certaines mesures. Donc, une image est utilisée comme requête pour rechercher les images contenant les mêmes objets ou décrivant la même scène que l'image requête. Pour cela, une base d'images ou le web sont utilisés comme espace de recherche [20].

II.4. Exigence et contraintes

Chaque cas d'utilisation exige des exigences à garantir pour bien fonctionner et des contraintes à respecter dans la reconnaissance de l'objet. Quelques-unes sont exposées dans ce qui suit :

- 1) **Temps d'évaluation (Evaluation Time) :** Dans le domaine industriel les données doivent être traitées en temps réel. Exemple : Le système de reconnaissance doit localiser les composants électroniques en une période de 10 à 50 ms pour assurer une très bonne vitesse de production. Le temps d'évaluation dépend de la taille de la zone (image) à inspecter.
- 2) **Précision :** Dans certain système la position de l'objet doit être localisée avec précision. Ex : Le système de reconnaissance doit localiser avec précision les composants électroniques, sinon ces composants seront abimer chaque fois où la position n'est pas parfaite.
- 3) **Fiabilité de reconnaissance :** Dans un système d'inspection, l'objectif principal est de réduire autant que possible le nombre de fausses rejets : bons objets classés comme défectueux et le nombre fausses acceptations : objets défectueux classés comme bons objets. Dans ce contexte le système de reconnaissance doit être fiable.
- 4) **Invariance :** Les algorithmes de reconnaissance doivent être insensibles à certaines variations des objets de la même classe dont ils ont la charge de reconnaître. Si cette variation n'existe pas alors le processus de reconnaissance sera parfait ! Cependant, le parfait n'existe pas, donc, l'algorithme doit être conçu de manière à maximiser la variance inter-classes et à minimiser la variance intra-classes.
- 5) **Luminosité :** L'apparence de l'objet diffère selon la variation de l'intensité, l'angle de projection et la couleur de la lumière utilisée. D'une manière générale, l'objet doit être reconnaît sont sensibilité aux variations de la lumière.
- 6) **Taille de la forme :** La surface de pixels occupée par l'objet dépend de sa distance du capteur d'acquisition de l'image, l'algorithme de reconnaissance d'objet doit être invariant au facteur de l'échelle.
- 7) **Rotation :** La disposition de l'objet à reconnaître dans une image n'est pas connue d'avance. Dans ce cas, l'algorithme doit prendre en considération toutes les positions (rotations) possibles que l'objet peut se procurer dans la scène.

- 8) Chevauchement avec l'arrière plan :** Les images ne contiennent pas que l'objet à reconnaître, mais toute une scène figure dedans constituant l'arrière plan de l'objet, ce dernier peut chevaucher avec l'objet (les données de l'arrière-plan sont dites corrélées avec ceux de l'objet). Donc, le système de reconnaissance ne doit pas être influencé par l'arrière plan.
- 9) Occultation d'une partie de l'objet :** Parfois la surface de l'objet n'est complètement affichée dans l'image, l'objet peut être partiellement masqué par d'autres objets dans l'image. L'objet doit être reconnaître même s'il est partiellement masqué.
- 10) Changement de l'angle de capture :** En général, l'acquisition d'une image se fait par projection d'une scène qui se produit dans un environnement 3D sur un plan 2D. Dans un plan 2D la projection des objets change selon l'angle sous laquelle la scène est capturée. L'insensibilité à l'angle de capture est une caractéristique désirable dans un système de reconnaissance, mais elle n'est pas toujours possible pour certaines formes. La conception du système de reconnaissance doit assurer un minimum d'invariance au point de capture au moins certaines pour formes. Seule la nature de l'application détermine les types d'exigences et contraintes sous lesquelles fonctionne le système de reconnaissance, quelques applications opèrent sur des images contenant beaucoup d'objets, donc, doivent gérer plusieurs contraintes en même temps : chevauchement avec l'arrière plan, changement de l'angle de capture, rotation, etc. Cependant, les applications industrielles opèrent sous des exigences très sévères limitant les contraintes du système de reconnaissance : souvent l'image contient un seul objet à reconnaître ou à inspecter, la lumière et l'angle de capture sont stable, etc. Ces exigences sont indispensables pour un fonctionnement en temps réel avec un taux d'erreur non significable [20].

II.5. Classification des méthodes de reconnaissance

La nature de chaque application, ses exigences et ses contraintes justifie le grand nombre d'approches et d'algorithmes de reconnaissance d'objets, il n'existe pas un schéma général applicable dans tout les cas pour reconnaître un objet, mais il existe beaucoup d'approches et de méthodes, chacune répond aux contraintes de l'application pour laquelle elle est proposée. Une classification de ces méthodes peut être faite selon différents critères : quelques-unes font référence aux propriétés des données représentant le modèle de l'objet, alors que d'autres font référence au schéma du processus de reconnaissance lui-même [20].

1) Représentation de l'objet

Principalement, il existe deux techniques pour représenter les données de l'objet à reconnaître : géométrique et apparence, les informations géométriques font référence aux limites de l'objet et sa surface : c-à-dire : contour et silhouette de l'objet. Les informations de contour sont souvent centré sur l'objet, les informations concernant la position des contours est affixé à un unique système de coordonnées de l'objet, et la création du modèle est souvent faite par un humain en utilisant les techniques de dessin CAD.

Au contraire de cela, les méthodes basées sur l'apparence utilise les caractéristiques de la région occupée par l'objet. Le modèle est créé à partir d'une phase d'apprentissage dans laquelle le système construit automatiquement un modèle de l'objet à reconnaître à partir d'une liste d'images contenant l'objet. Dans ce cas le modèle est centré sur les conditions de capture des images utilisées dans l'apprentissage [20].

2) Domaine des données de l'objet

Les données du modèle peuvent référées à des propriétés locales de l'objet (ex : certains angle de l'objet) ou des caractéristiques générales de l'objet (ex : surface, périmètre, centre d'inertie). Dans le cas de données local, le modèle est un ensemble de données caractérisant différents parties de la surface occupée par l'objet. Dans le cas d'une représentation globale différents caractéristiques sont utilisées dans un vecteur d'attributs. Cette représentation est propice pour des objets simple (cercles, rectangle, ... en 2D ou cylindre, battons, ... en 3D).

Il est important de noter que les données locales sont convenables spécialement pour des objets complexes et structurés. Surtout dans le monde industriel ou les pièces peuvent être décrite par des lignes, des cercles des angles des arcs, etc. ces primitives peuvent être modélisés est cherchées localement. L'utilisation de données local permet de faire l'invariance par apport à l'occultation, car chaque caractéristique est détectée séparément ; si quelques unes manquent, le reste permet de reconnaître l'objet.

3) Variations possibles de l'objet

L'autre critère est la variation qui peut se trouver entre les objets de la même classe. Dans le domaine industriel, il y a une minime variance entre les éléments de la même classe, c'est pourquoi un modèle rigide peut être appliqué.

4) Qualité des données de l'image

La qualité de l'image a un important impact sur l'algorithme de reconnaissance. Dans le domaine industriel, la qualité est supérieure, pas de variation de la lumière, pas d'arrière plan, pas d'autres objets, un seul angle de capture, etc. mais dans un système de surveillance par exemple : la lumière change, la position de capture change (caméra commandée à distance), les objets bougent donc changent de rotation, de position, etc.

5) Le processus de reconnaissance

Pour reconnaître un objet, une mesure de correspondance est effectuée dans un certain niveau de l'algorithme, le modèle de l'objet ou une partie de lui doit être comparé au contenu de l'image pour mesurer la similitude/dissimilitude entre l'objet et le contenu de l'image. Quelques algorithmes essaient d'optimiser les paramètres de transformations caractérisant la relation entre le modèle et sa projection sur le plan de l'image pour cela une affine transformation est utilisée. D'autres approches essaient de trouver une correspondance entre les attributs caractéristiques de l'objet et ceux qui sont extraits de l'image.

6) Domaine des éléments de données utilisés dans la mesure de correspondance

Typiquement, les données utilisées dans la mesure de correspondance entre le modèle et le contenu de l'image peuvent être divisées en trois catégories : ligne de valeurs d'intensité des pixels, attributs de bas niveau comme le contour, et des attributs de haut niveau comme des lignes, des arcs, etc.

2. Conclusion

Dans ce chapitre, nous avons défini que la reconnaissance de forme est une science d'élaboration d'algorithmes permettant de classer des objets dont l'aspect varie par rapport à un objet type. Son objective est la conception et l'analyse des algorithmes de classification applicables aux présentations d'images, de sons ou de signaux de toute nature. Ensuite, nous avons abordé : les différents rôles, le schéma général, les approches, les méthodes, et les différents champs d'application de la reconnaissance des formes. Aussi nous avons présenté les différents algorithmes de reconnaissance d'objet, les domaines d'applications, exigences et contraintes, et finalement nous avons terminé par une classification des méthodes de reconnaissance.

Chapitre IV : Conception et implémentation

- 1- Introduction
- 2- Conception
- 3- Implémentation de système
- 4- Etude comparative
- 5- Conclusion

Chapitre IV

Conception et implémentation du système : ACLB « Automatic Construction of Learning Base »

1. Introduction

Une base d'apprentissage est un ensemble d'attributs caractéristiques discriminants qui sont utilisées afin de limiter la quantité d'informations utilisée dans l'apprentissage et améliorer leur qualité. Le système ACLB est un système qui permet la construction automatique d'une base d'apprentissage, à partir d'un ensemble d'images collectées du web selon des critères bien définis. Les résultats retournés par le système seront filtrés pour ne garder que les images pertinentes qui se rapportent au sujet recherché. Ces derniers seront alors utilisés pour construire la base d'apprentissage.

Ce chapitre est composé de deux parties : la première permet de présenter la conception du système ACLB et la deuxième montre l'implémentation du système.

2. Conception

2.1. Objectifs du système

Les objectifs assurés par notre système « ACLB » sont :

- **Collecte d'un ensemble d'images à partir du web** : cette partie consiste à chercher automatiquement du web un ensemble d'images selon une requête posée par l'utilisateur. Un dictionnaire dans le moteur de recherche est utilisé pour enrichir automatiquement les différentes requêtes posées.
- **Restriction de l'ensemble d'images** : afin de restreindre l'ensemble des images collectées, et ne garder que celles qui sont vraiment en rapport avec les critères recherchés, des méthodes de recherche d'images par contenu sont utilisées (l'histogramme, la texture et la combinaison des deux).
- **Filtrage à base d'objet choisi** : dans la recherche d'images par contenu, les images utilisées pour la recherche peuvent contenir des objets autres que l'objet recherché et qui peuvent influencer d'une façon négative la recherche. Pour pallier à ce problème, nous avons utilisé la méthode de corrélation et la méthode de corrélation à variante du Sobel qui permettent de détecter un objet dans image.

- **Evaluation** : puisque plusieurs méthodes de recherche ont été réalisées, nous avons sélectionné quatre critères permettant l'évaluation de ces méthodes. Cette évaluation facilite à l'utilisateur le choix de la méthode à appliquer pour construire sa base d'apprentissage.
- **Construction de la base d'apprentissage** : à la fin du collecte d'images, nous procédons à l'extraction d'attributs caractéristiques pour limiter la quantité d'informations utilisée dans l'apprentissage et améliorer la qualité d'apprentissage en ne gardant que les attributs caractéristiques discriminants et éliminant ceux redondants ainsi que le bruit. Ces attributs caractéristiques sont utilisés pour construire la base d'apprentissage.

2.2. Architecture du système

L'architecture du système implémenté est la suivante :

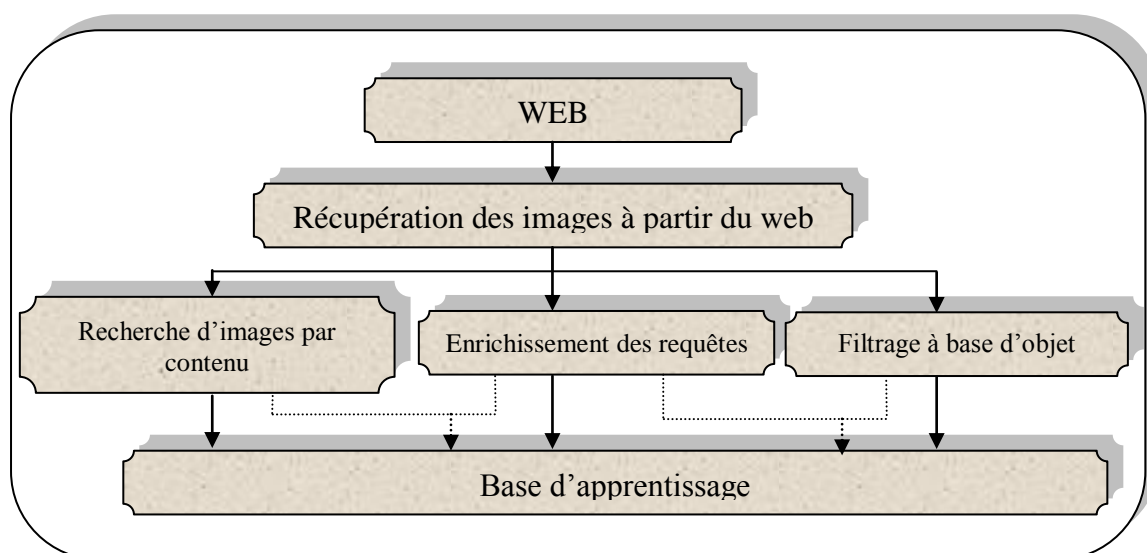


Figure IV.1 : Étapes de construction de la base d'apprentissage.

2.2.1. Web

Le moteur de recherche utilisé pour la recherche et la récupération des images à partir du web est le moteur « LYCOS ». (Voir *figure IV.2*)



Figure IV.2 : Présentation du site web utilisée par notre système « LYCOS »

2.2.2. Récupération des images à partir du web

On récupère les 96 premiers résultats retournés par le moteur « Lycos », puis on les stocke sur disque dans le dossier portant le même nom que celui de la requête entrée par l'utilisateur lors de la recherche. (Voir *figure IV.3*).

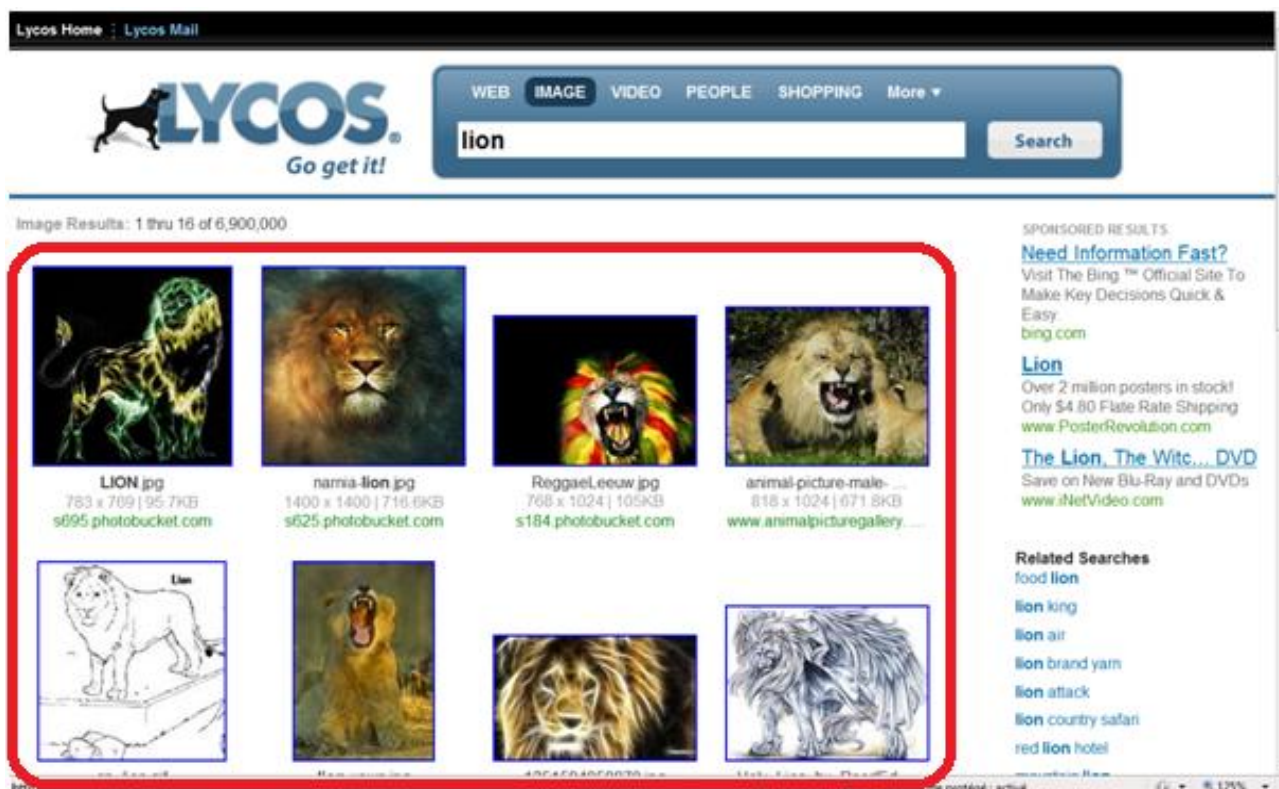


Figure IV.3 : Présentation des résultats de recherche sur « LYCOS »

2.2.3. Enrichissement des requêtes

Notre système offre un dictionnaire pour enrichir automatiquement les différentes requêtes posées par l'utilisateur, ce dernier est extrait à partir « Related Searches » présentées sur le moteur « Lycos ». (Voir *figure IV.4*).

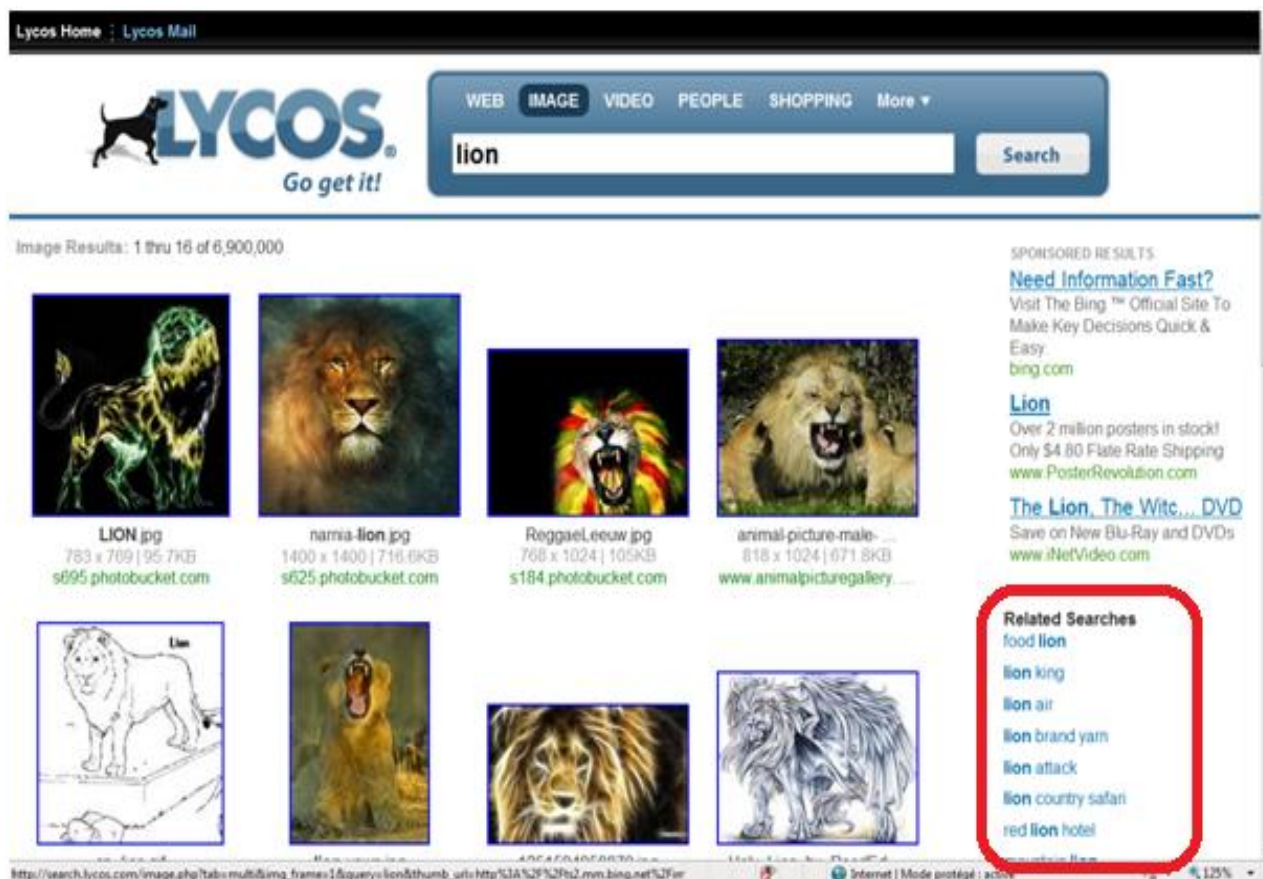


Figure IV.4 : Présentation des recherches liées à la requête sur « LYCOS »

2.2.4. Recherche d'images par contenu

On implémenté trois méthodes de recherche par le contenu qui sont :

2.2.4.1. Recherche d'image en utilisant la couleur

L'utilisation d'histogrammes est une technique très utilisée pour la couleur (voir *Figure IV.5*). Avec cette méthode, on doit d'abord calculer l'histogramme d'image (l'histogramme d'une image peut être présenté par un vecteur dont chaque composant est un nombre de pixels de couleur correspondant à son indice). Cette méthode est présentée par la *Figure IV.5*.

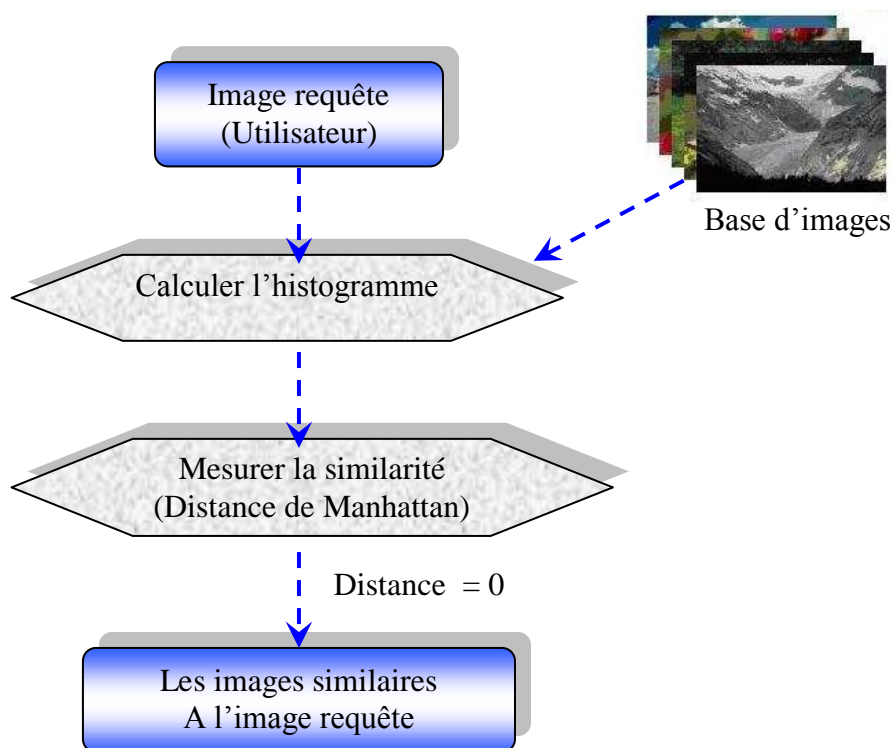


Figure IV.5 : Les étapes de recherche d'image en se basant sur l'histogramme.

1) Calcule d'histogramme

Un histogramme représente le nombre de points de l'image en fonction des différents niveaux de gris. Si on prend par exemple une image 8 bits (u), on peut calculer la présence d'un niveau de gris v dans l'image par :

$$\forall v \in \{0, \dots, 255\} \quad h_u(v) = \text{Nombre de pixels d'intensité } v \quad (\text{IV.1})$$

2) Calcul de similarité entre les images

On compare l'histogramme de l'image requête aux histogrammes des images de la base. Pour cela les histogrammes sont considérés comme des vecteurs, et on peut alors calculer une distance entre les deux vecteurs. La distance utilisée est la distance Manhattan :

$$d(U, V) = \sum_{i=0}^{n-1} |U_i - V_i| \quad (\text{IV.2})$$

2.2.4.2. Recherche d'image en utilisant la texture

En 1973, Haralick [21] a proposé une méthode se basant sur les matrices de cooccurrences de niveaux de gris. Elle est l'une des méthodes les plus célèbres pour analyser la texture (voir *Figure IV.6*).

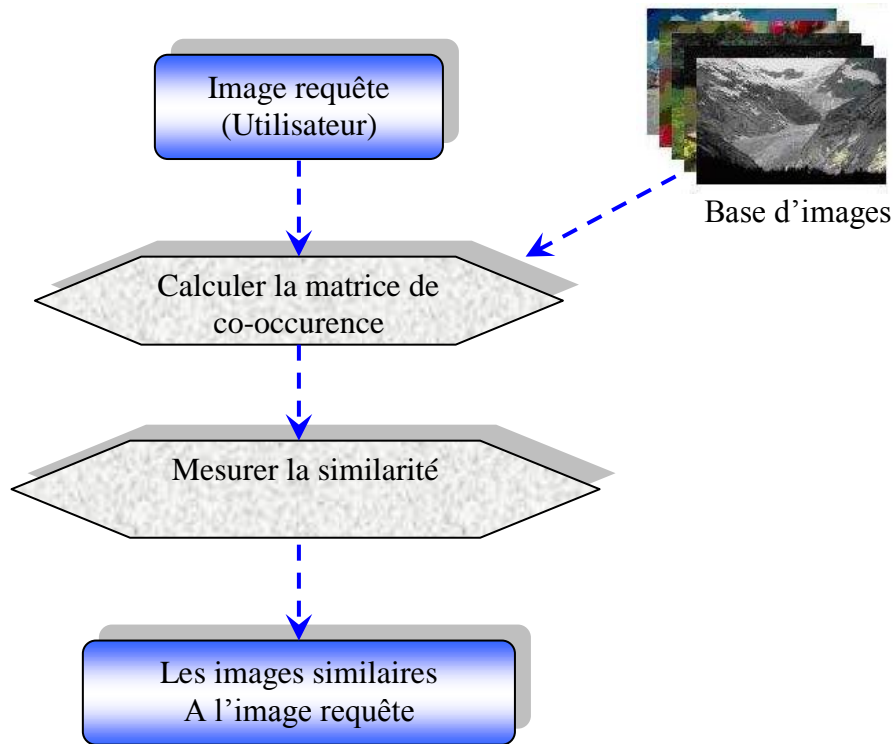


Figure IV.6 : Les étapes de recherche d'image en se basant sur la texture.

Cette méthode permet de déterminer la fréquence d'apparition d'un "motif" formé de deux pixels séparés par une certaine distance d dans une direction particulière par rapport à l'horizontale. Afin de limiter le temps de calcul, on prend généralement comme valeurs 0° , 45° , 90° , 135° , 180° et 1 pour la valeur de d (voir **figure IV.7**).

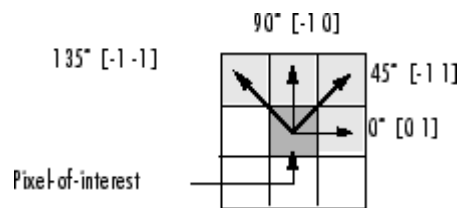


Figure IV.7 : Présentation des distances et des directions.

La **Figure IV.8** montre un exemple de calcul des P_{ij} (où P est la matrice de cooccurrence) à partir d'une petite image 4×5 composée de huit niveaux de gris (1, ..., 8). Cet exemple se limite au cas $d=1$, l'élément $(1,1)$ de la matrice P (1,0) est égal à 1 cela signifie qu'il existe une seule configuration dans l'image où un pixel de niveau de gris 1 est séparé horizontalement d'un autre pixel de niveau de gris 1 par distance 1 dans la direction 0° .

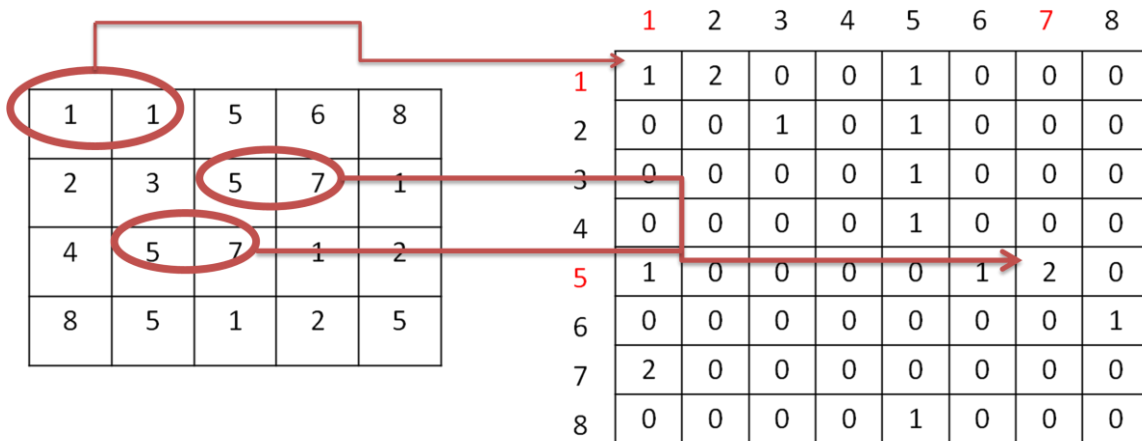


Figure IV.8 : Exemple de matrices de cooccurrences construites à partir d’une image 4x5 composées de huit niveaux de gris.

Afin d’estimer la similarité entre les matrices de cooccurrences, Haralick [21] a proposé 14 caractéristiques statistiques extraites à partir de cette matrice. Nous avons utilisé quatre parmi ces derniers : l’inertie (T1), l’entropie (T2), l’énergie(T3), l’homogénéité (T4).

La distance entre deux images H, I se calcule par la formule suivante :

$$D_{texture}(H, I) = \sqrt{(T_{H1} - T_{I1})^2 + (T_{H2} - T_{I2})^2 + (T_{H3} - T_{I3})^2 + (T_{H4} - T_{I4})^2} \quad (IV.3)$$

Où :

- T_{H1} : L’énergie de H. T_{I1} : L’énergie de I.
- T_{H2} : L’entropie de H. T_{I2} : L’entropie de I.
- T_{H3} : Le contraste de H. T_{I3} : Le contraste de I.
- T_{H4} : L’homogénéité de H. T_{I4} : L’homogénéité de I.

2.2.5. Filtrage à base d’objet choisi (méthode de corrélation)

La méthode de corrélation croisée (cross correlation) est appliquée entre le contenu de l’image et les données du modèle. Dont ce dernier est une image template qui est un prototype du niveau de gris de l’objet choisi par l’utilisateur. La corrélation 2D est un exemple d’approche basée sur l’apparence, car le modèle dépend exclusivement sur l’apparence de la surface occupée par le modèle dans les images de la base.

2.2.5.1. Corrélation

La corrélation est mesurée entre le modèle et les images de la base pixel par pixel d’une surface égale à celle du modèle dans l’image pour essayer toutes les positions possibles dans l’image. Et un coefficient de corrélation croisée normalisée (Normalized cross Correlation Coefficient (NCC)) est calculé à chaque mouvement (voir la *figure IV.9*).

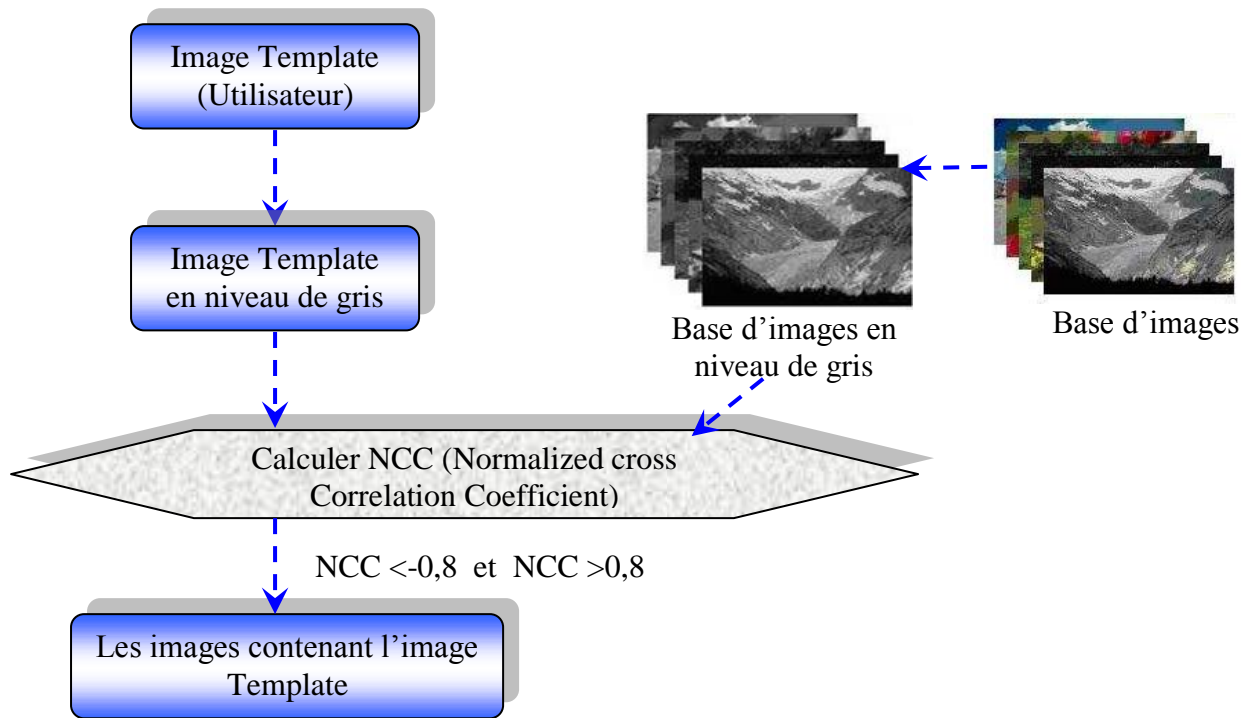


Figure IV.9 : Les étapes de la méthode de corrélation.

$$NCC(a, b) = \frac{\sum_{x=0}^W \sum_{y=0}^H (I_s(x+a, y+b) - \bar{I}_s) * (I_T(x, y) - \bar{I}_T)}{\sqrt{\sum_{x=0}^W \sum_{y=0}^H (I_s(x+a, y+b) - \bar{I}_s)^2 * \sum_{x=0}^W \sum_{y=0}^H (I_T(x, y) - \bar{I}_T)^2}}$$

(IV.4)

Où $NCC(a, b)$ est le coefficient de corrélation croisée normalisée (Normalized cross Correlation Coefficient (NCC)) au déplacement $[a, b]$ entre l'image et le modèle de la scène. $I_s(x, y)$ et $I_T(x, y)$ désignent respectivement l'intensité de l'image de la scène et le modèle à la position $[x, y]$, \bar{I}_s , \bar{I}_T leurs moyennes, W , H est la largeur et la hauteur de l'image modèle. Parce que le dénominateur sert de normalisation NCC terme peut varier entre -1 à 1.

- Les valeurs positives élevées indiquent que l'image et le modèle de la scène sont très similaires.
- Une valeur de 0 indique que leur contenu n'est pas corrélé.
- Les valeurs négatives sont la preuve de l'inverse contenu [22].

La *figure IV.10* illustre le mouvement d'un modèle (vert) à travers une image exemple au cours du processus de calcul d'une corrélation.

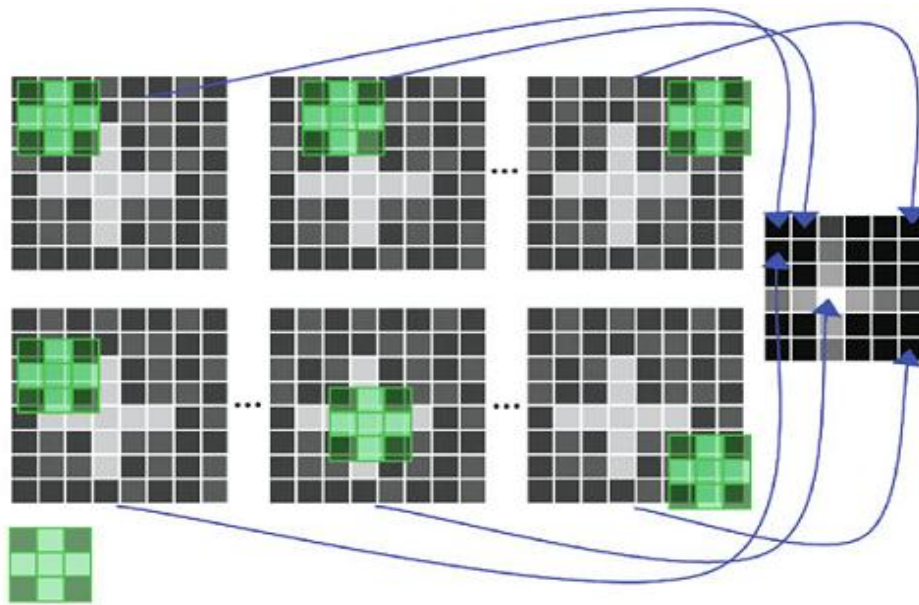


Figure IV.10 : Exemple d'application de la méthode de corrélation.

2.2.5.2. Corrélation (variante Sobel)

Au lieu d'utiliser l'image d'intensité originale telle quelle est pour la mesure de corrélation (ici on parle de la template sous forme de niveau de gris), il est possible d'utiliser une image gradient : les images de base sont prétraitées avec un filtre de contour ce qui donne une template prétraitée, donc la corrélation est calculée entre les images filtrées (Figure IV.11).

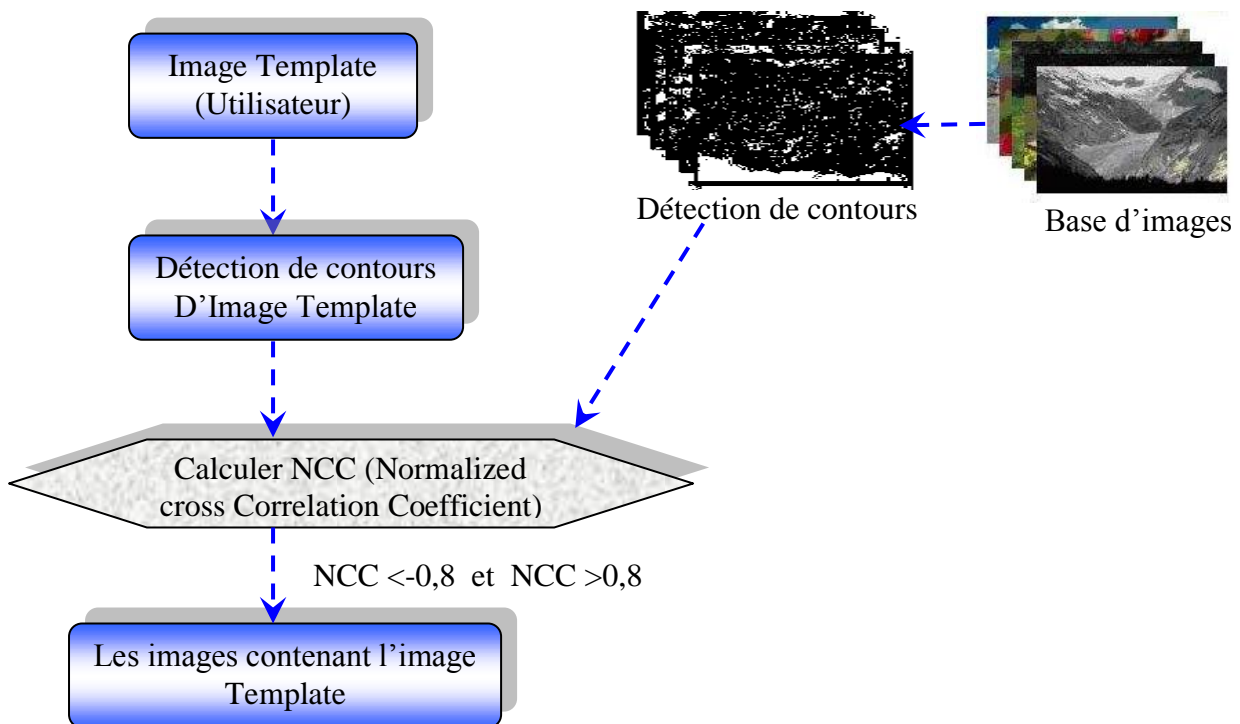


Figure IV.11 : Les étapes de la méthode de la corrélation (variante Sobel).

- Détection des contours (*Figure IV.12*)

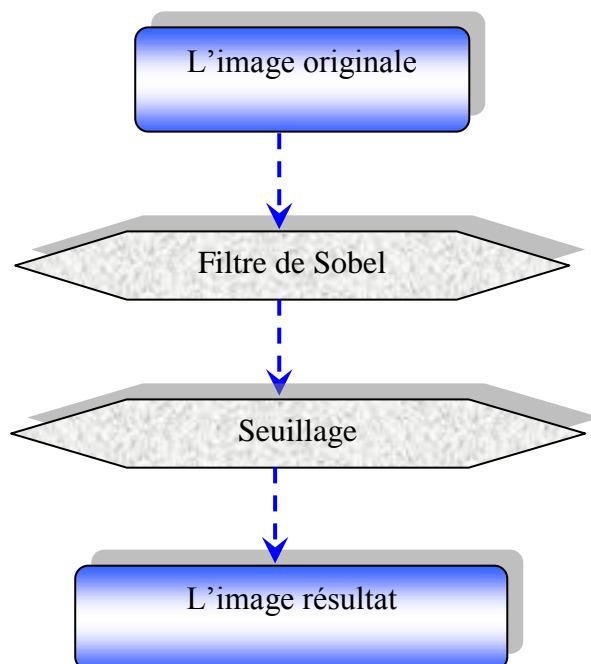


Figure IV.12 : Les étapes de la détection des contours.

A) Le filtre de Sobel

Le filtre de Sobel détecte séparément les bords horizontaux et verticaux sur une image en niveaux de gris. Les images en couleurs sont transformées du RVB en niveaux de gris. L'application du filtre aboutit à une image transparente avec de fines lignes noires et quelques restes de couleurs [Web 5].

L'opérateur utilise des matrices de convolution. La matrice (généralement de taille 3×3) subit une convolution avec l'image pour calculer des approximations des dérivées horizontale et verticale. Soit A l'image source, G_x et G_y deux images qui en chaque point contiennent des approximations respectivement de la dérivée horizontale et verticale de chaque point. Ces images sont calculées comme suit [Web 6]:

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A \text{ et } G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (\text{IV.5})$$

En chaque point, les approximations des gradients horizontaux et verticaux peuvent être combinées comme suit pour obtenir une approximation de la norme du gradient:

$$G = \sqrt{G_x^2 + G_y^2} \quad (\text{IV.6})$$

B) Seuillage

L'opération dite de "seuillage simple" consiste à mettre à zéro tous les pixels ayant un niveau de gris inférieur à une certaine valeur (appelée *seuil*, en anglais *threshold*) et à la valeur maximale les pixels ayant une valeur supérieure. Ainsi le résultat du seuillage est une image binaire contenant des pixels noirs et blancs, c'est la raison pour laquelle le terme de *binarisation* est parfois employé. Le seuillage permet de mettre en évidence des formes ou des objets dans une image. Toutefois la difficulté réside dans le choix du seuil à adopter [Web 7].

$$g(x, y) = \begin{cases} 0 & \text{si } f(x, y) \leq T \\ 1 & \text{si } f(x, y) \geq T \end{cases} \quad (\text{IV.7})$$

Avec :

$$T = T[(x, y), p(x, y), f]$$

g : Image binaire

(x, y) : Coordonnées du pixel

$g(x, y)$: Propriété locale au pixel

f : Image originale

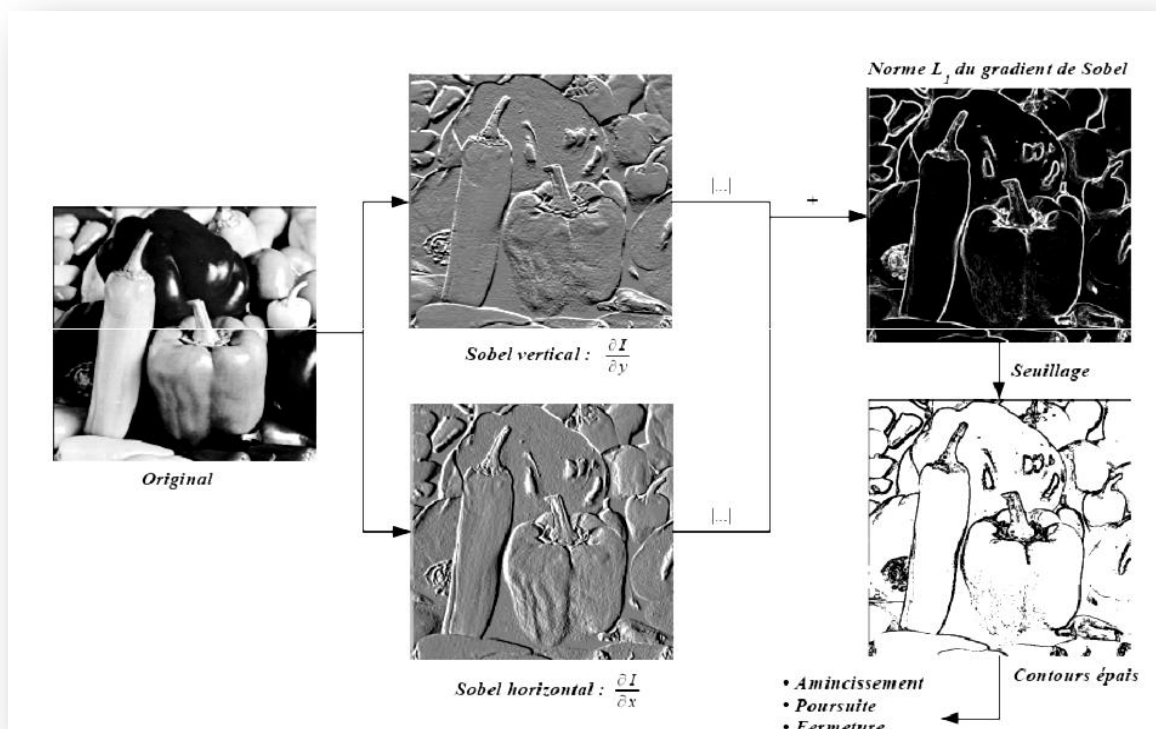


Figure IV.13 : méthode de détection des contours en utilisant le filtre de Sobel.

3. Implémentation du système

Dans cette partie nous allons aborder l'aspect implémentation de l'application réalisée « **Construction Automatique des Bases d'Apprentissage** ». Parler de l'implémentation revient à détailler l'aspect matériel, l'environnement de développement et les différents modules qui composent le logiciel.

3.1. Aspect matériel

Notre projet a été développé sur un pc :

- ✚ Type : PC portable HP 550.
- ✚ Processeur : Intel Pentium Core 2 Duo (1.8 GHZ).
- ✚ RAM : 2 Go.
- ✚ Disque dure : 160 Go.

3.2. Environnement de développement

L'application réalisée a été implémentée avec le langage **JAVA version 6** sous l'environnement **Eclipse HELIOS (version 2000-2011)**.

3.2.1. Présentation de JAVA

Java est un langage de programmation développé par Sun Microsystems, il n'a que quelques années de vie (les premières versions datent de 1995). Le système Java comporte plusieurs parties : un environnement, le langage, les interfaces de programmation d'application, et diverses bibliothèques de classes.

Malgré la nouveauté de Java, il a réussi à intéresser et intriguer beaucoup de développeurs à travers le monde et il a été développé dans le but d'augmenter la productivité des programmeurs [23]. Sa réputation est due aux principaux avantages suivants :

- C'est un langage orienté objet dérivé du C, mais plus simple que C.
- Il est multiplateforme : tous vos programmes tournent sans modification sur toutes les plateformes où existe Java.
- Il est doté en standard d'une riche bibliothèque de classes, comprenant la gestion des interfaces graphiques (fenêtres, boîtes de dialogue, contrôles, menus, graphisme...), la programmation multithreads (multitâches), la gestion des exceptions, les accès aux fichiers et au réseau (notamment Internet)... etc. [24].

3.2.2. Pour quoi Java ?

Contrairement à d'autre langage, la simplicité de java est en fait directement liée à la simplicité du problème à résoudre. Ainsi, un problème de définition de l'interface utilisateur peut être un véritable casse-tête avec certains langages, avec java, créer des fenêtres, des boutons et des menus est toujours d'une simplicité extrême, ce qui laisse au programmeur la possibilité de se concentrer sur son problème à résoudre.

- Java est un environnement complet qui offre des flexibilités et des possibilités importantes
- Java est un langage relativement simple à apprendre et lire.
- Java offre une probabilité grâce à la machine virtuelle (JVM) [25].

3.2.3. Présentation d'Eclipse

Eclipse est un environnement de développement intégré (Integrated Development Environment) dont le but est de fournir une plate forme modulaire pour permettre de réaliser des développements informatiques (voir *Figure IV.14*).

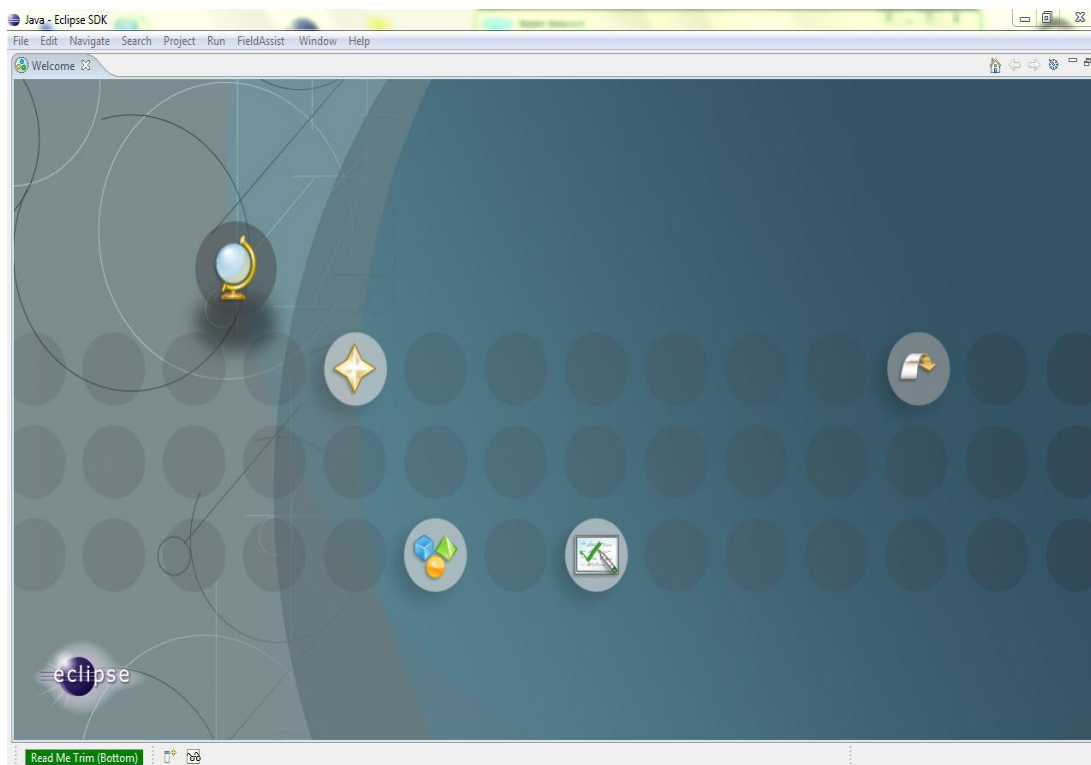


Figure IV.14 : Environnement Eclipse.

3.2.4. Représentation de l'image dans Java

L'acquisition ou la numérisation d'une image procède d'abord par le chargement d'une image sur un tampon (Buffered), en suite transférée sur un tableau ($S[i]$ / i est le nombre de pixel) pour la préparation de traitement.

- **Cas d'image niveau de gris** : en chaque point $s = i$ de S de l'image, on enregistre la luminosité $Y(s)$ en ce point. $Y(s)$ représente alors le niveau de gris en s ($Y(s)$ est de valeur entière allant de 0 si le point est un noir complet, à 255 si au contraire il est du blanc saturé). Une image numérique est mathématiquement une application :

$$S = \{1 \dots, M * N\} \text{ dans } \{0, \dots, 255\}; s \text{ qui donne } Y(s)$$

- **Cas d'une image couleur** : Pour une image en couleur on mesure la luminosité en un point s selon les trois canaux rouge, vert et bleu (RVB).

L'image couleur est alors représentée par un tableau de $N * M$ case où chaque case est de la forme hexadécimale suivante (*Figure IV.15*) [25]:

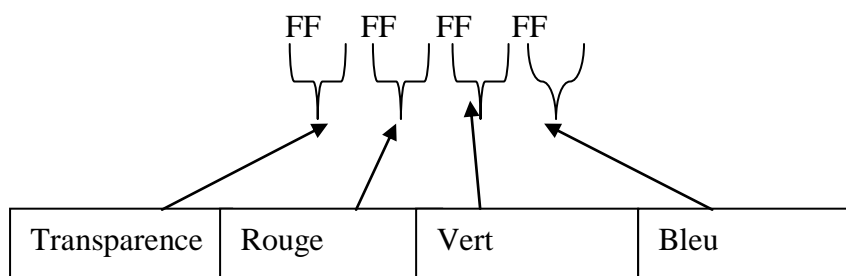


Figure IV.15 : Représentation d'un pixel couleur.

3.3. La réalisation de l'application

Le langage JAVA est basé sur la manipulation des packages et des classes et dans notre logiciel on a utilisé deux packages :

- **package « resultat_web »** : pour la manipulation des codes sources de la page web en HTML (ex : extraction des balises).
- **package « traitement »** : pour récupérer les images à partir des web puis effectuer des traitements pour garder qu'une séquence d'images pertinentes « base d'apprentissage ».

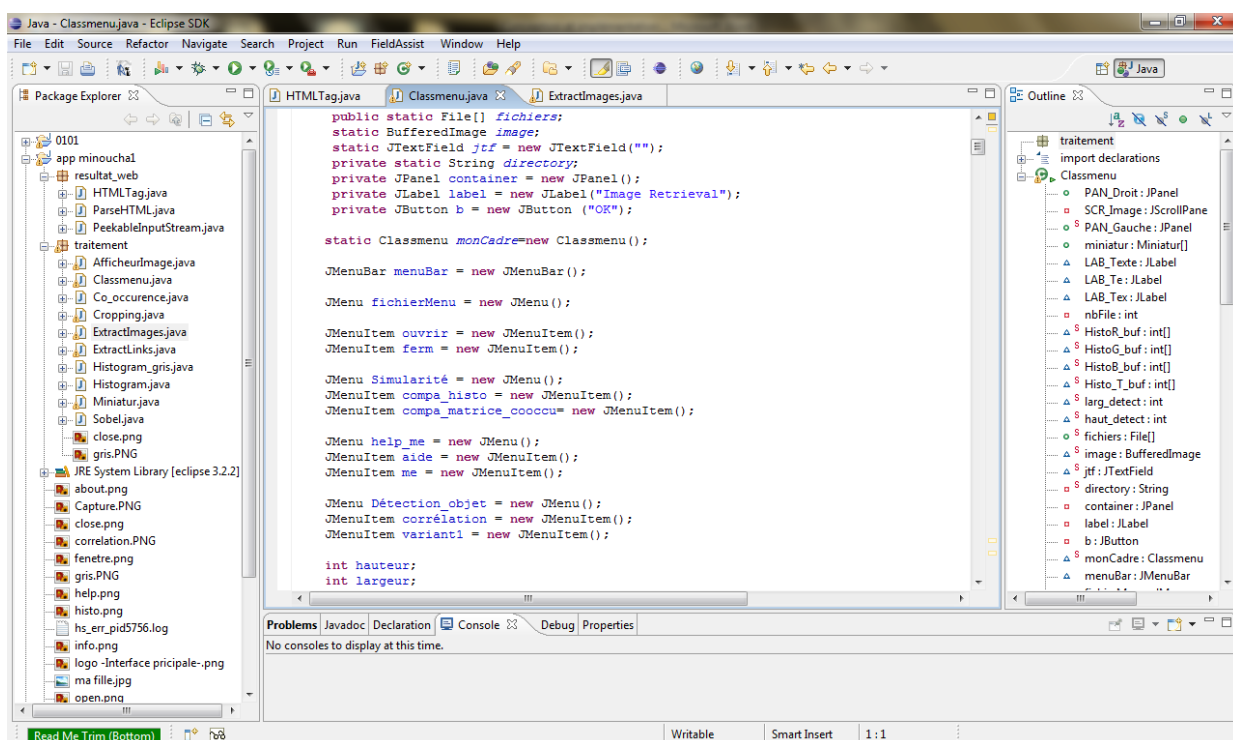


Figure IV.16 : Exemple extrait du code.

A) Package « resultat_web »

Il se compose de 3 classes :

- **ParseHTML.java** : C'est la classe qui analyse le HTML et génère comme sortie les textes bruts des balises HTML.
- **PeekableInputStream.java** : Il s'agit d'un flux d'entrée spéciale permet au programme de prendre un ou plusieurs caractères à venir dans le fichier.
- **HtmlTag.java** : Cette classe contient une seule balise HTML. Cela permet à la classe HtmlTag de tenir une collection d'attributs, tout comme une balise HTML.

B) package « traitement »

Dans ce package on utilisé 16 classes et parmi ces classes on cite :

- **Launch_application.java** : assure le lancement de l'application et permet l'accès à ses différentes fonctions.
- **ExtractImages.java** : récupère les images à partir du web tout en tenant compte du mot clé saisi par l'utilisateur.
- **ExtractLinks.java** : enrichit la requête en récupérant les recherches reliées au mot clé saisi par l'utilisateur.
- **Histogram.java** : calcule l'histogramme de l'image.
- **Co_occurrence.java** : calcule la matrice de cooccurrence de l'image.

- **Histogram_gris** : transforme l'image en niveau de gris puis calcule son histogramme.
- **Sobel.java** : détecte les contours pour appliquer la méthode de corrélation (variante Sobel).
- **Cropping.java** : sélectionne une partie à partir de l'image requête choisie par l'utilisateur.
- **P_R_P.java** : calcule les mesures d'évaluation du système.
- **Evaluation.java** : évalue les images de la base (si elles sont similaires à l'image requête ou non).

3.4. Présentation de l'application

Au démarrage de l'application, la page suivante est affichée lors de lancement de l'exécutable (*figure IV.17*).

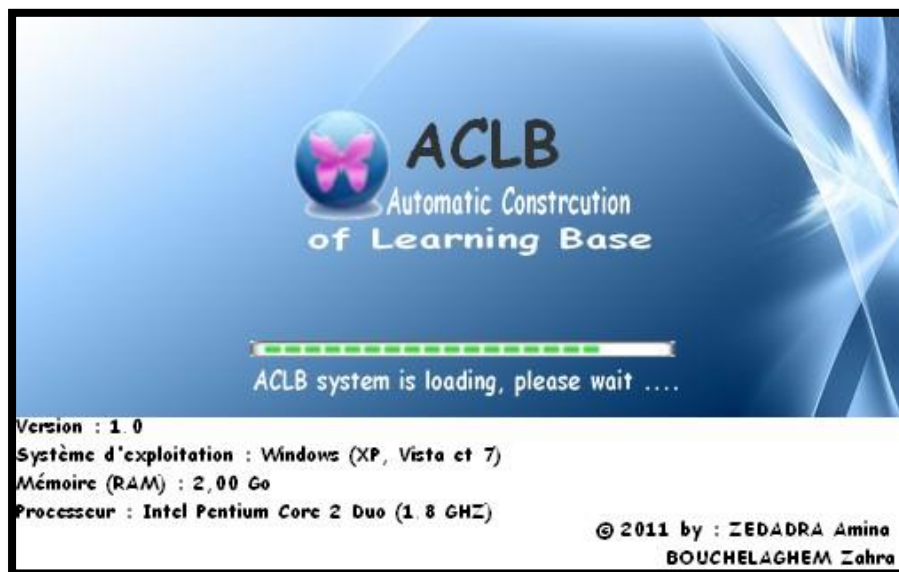


Figure IV.17 : Démarrage de l'application.

3.4.1. Fenêtre principale

Après la fin de chargement de la fenêtre de démarrage la fenêtre principale s'affichera :

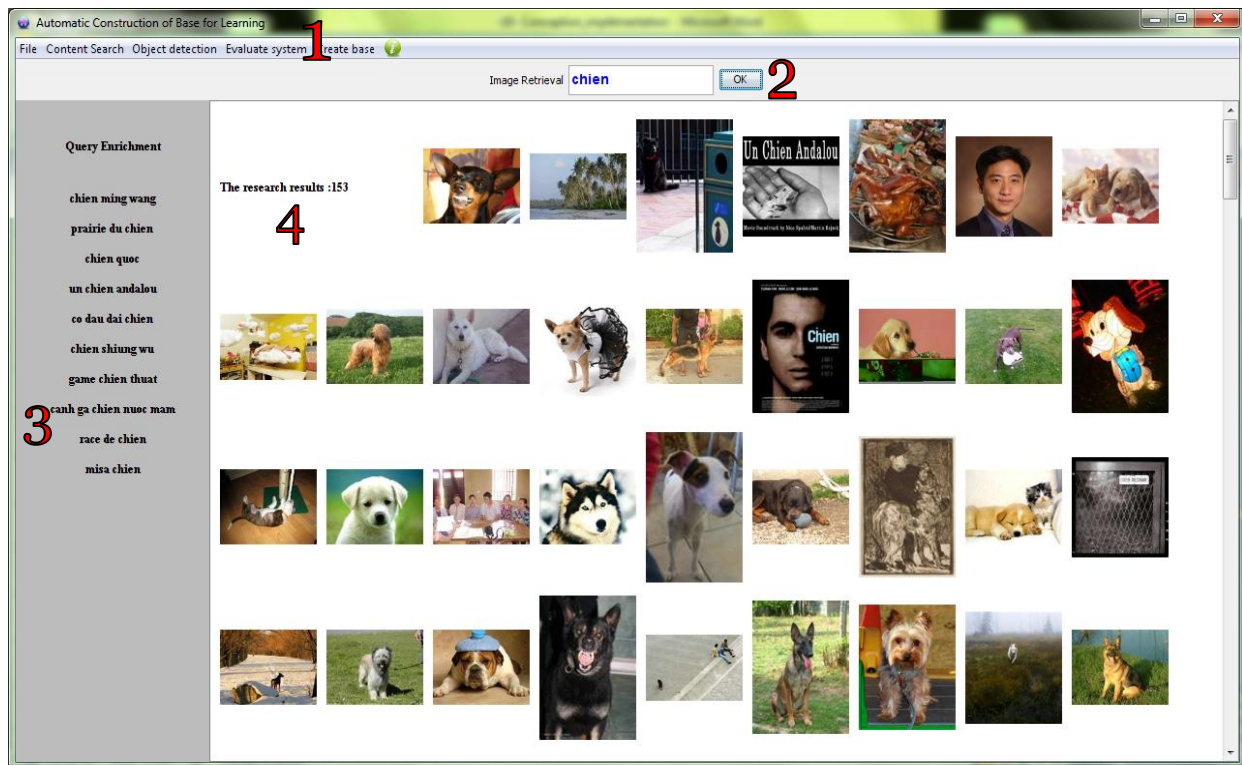


Figure IV.18 : Fenêtre principale de « ACLB ».

- 1 - Menu de l'application.
- 2 - La zone de texte dont laquelle l'utilisateur saisie sa requête « mot clé ».
- 3 - Affichage du dictionnaire du web qui est utilisé pour enrichir automatiquement les différentes requêtes posées par l'utilisateur.
- 4 - Affichage des résultats de recherche.

3.4.2. Chargement d'image requête

La première opération à effectuer après la récupération des résultats (les images), est de charger une image requête au choix d'utilisateur.

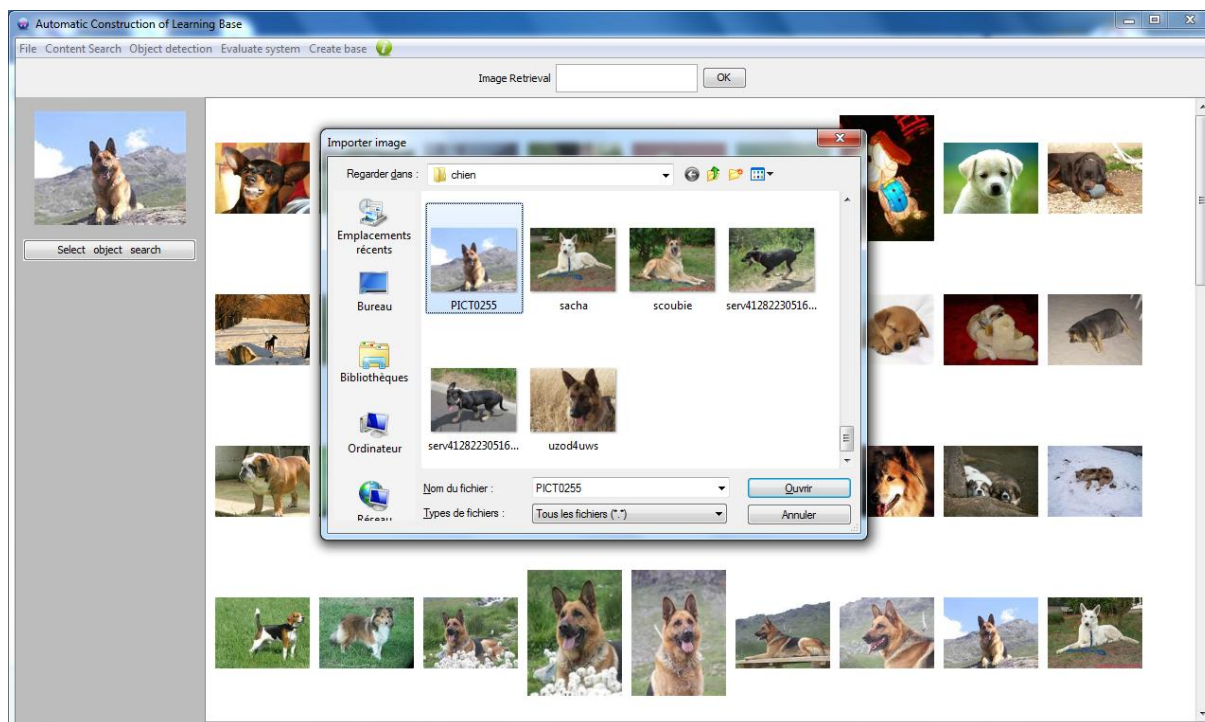


Figure IV.19 : Chargement de l'image requête.

3.4.3. Méthode de recherche par la couleur

En utilisant l'histogramme de l'image requête, on calcule la similarité entre les images récupérées du web on ne garde que celles qui sont vraiment en rapport avec l'image requête.

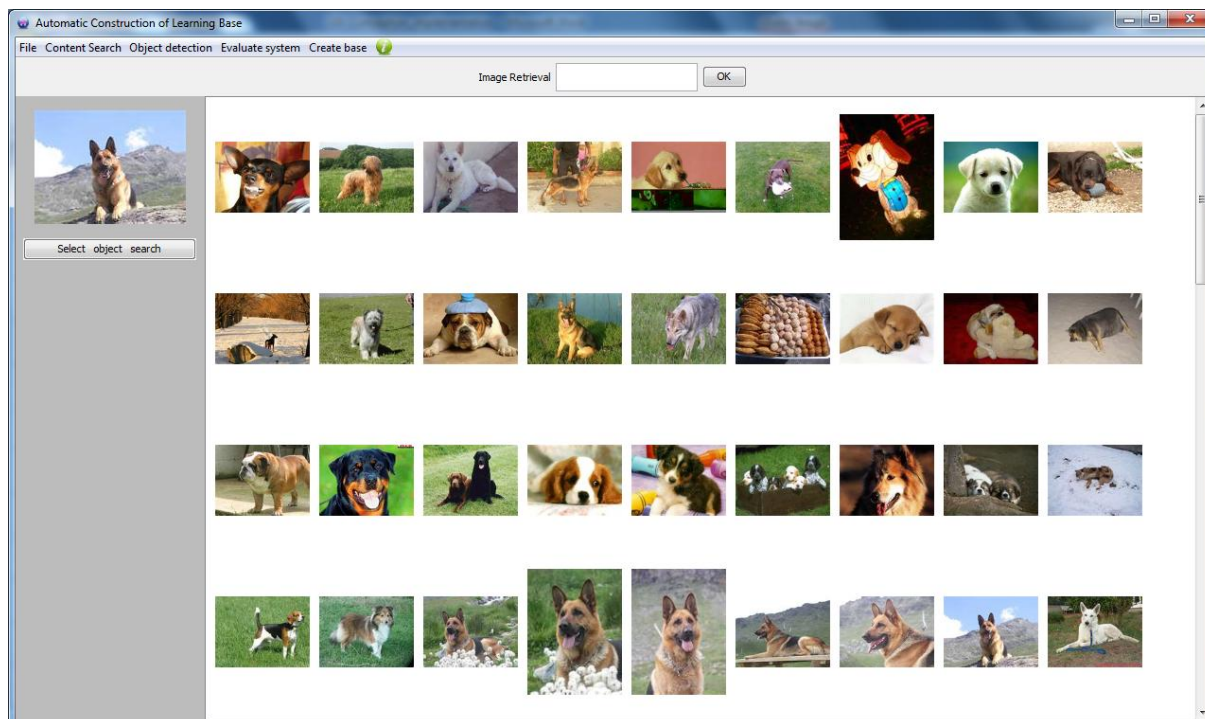


Figure IV.20 : Méthode de recherche par la couleur.

3.4.4. Méthode de recherche par la texture

En utilisant la matrice de cooccurrence calculée à partir de l'image requête (en niveau de gris), on calcule la similarité entre les images récupérées du web et on ne garde que celles qui sont vraiment en rapport avec l'image requête.

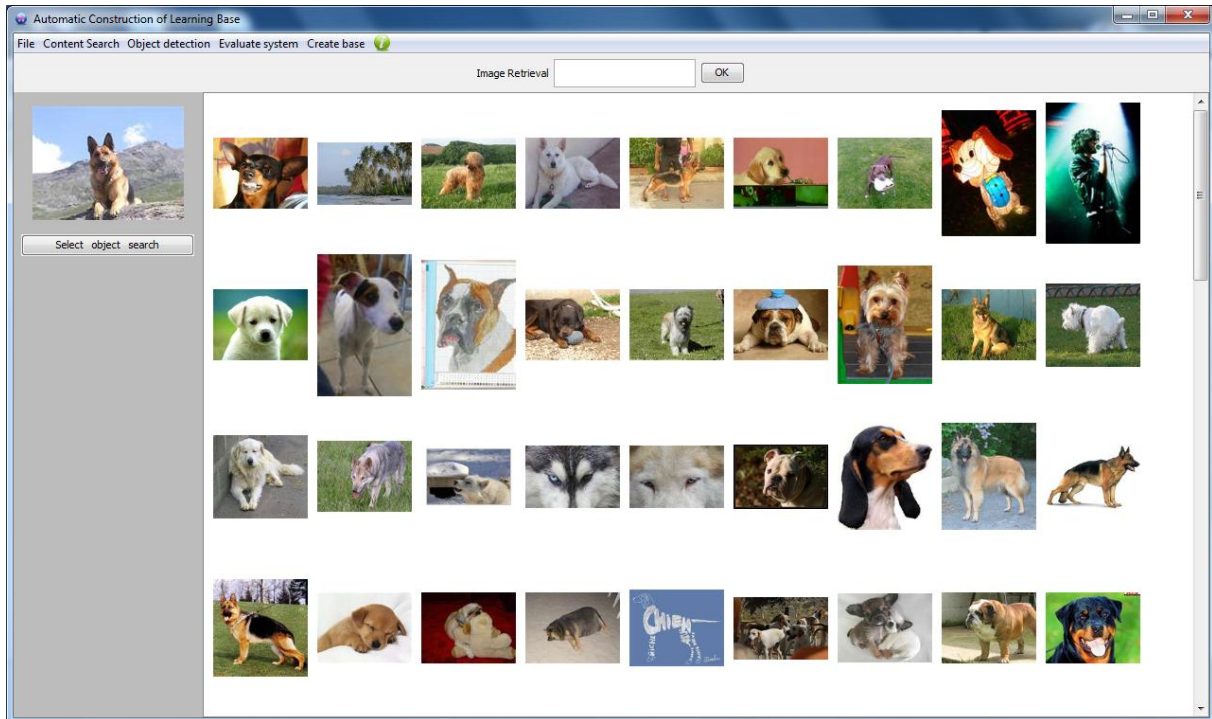


Figure IV.21 : Méthode de recherche par la texture.

3.4.5. Combinaison des deux méthodes (couleur et la texture)

En utilisant l'histogramme et la matrice de cooccurrence calculée à partir de l'image requête (en niveau de gris), on calcule la similarité entre les images récupérées du web et on ne garde que celles qui sont vraiment en rapport avec l'image.

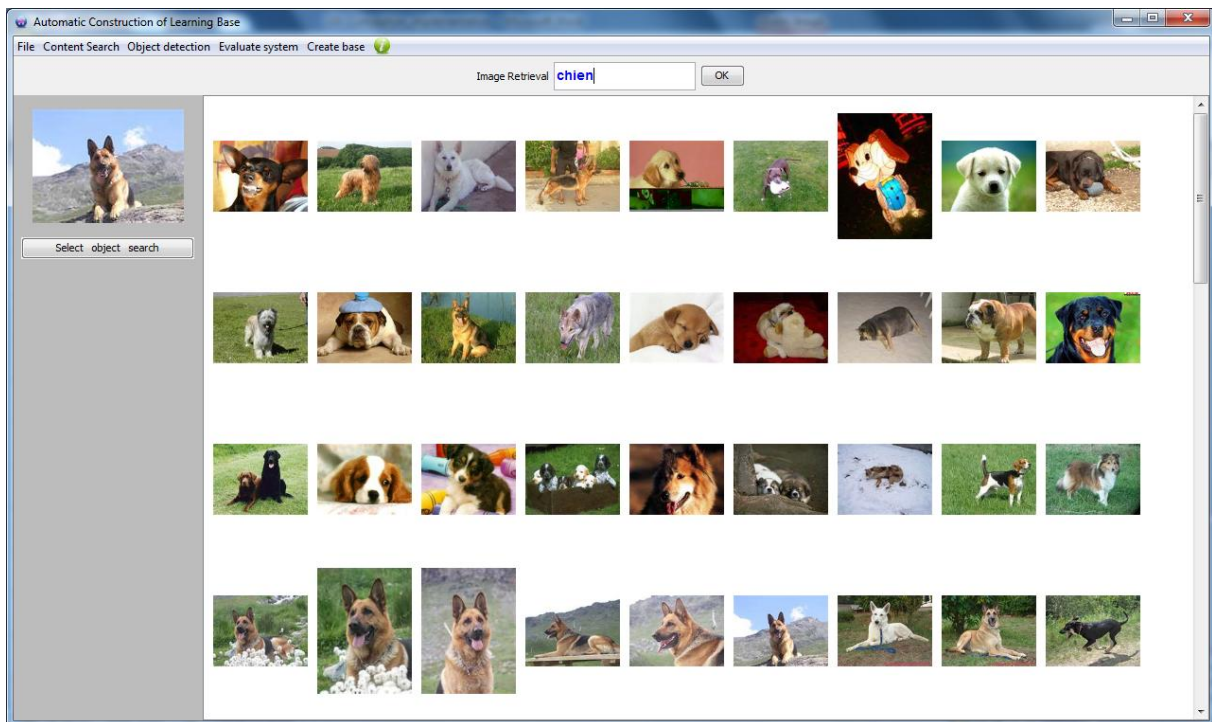


Figure IV.22 : Combinaison des deux méthodes (couleur et la texture).

3.4.6. La méthode de la corrélation

Dans cette méthode on cherche à laisser que les images qui contiennent l'objet (en niveau de gris) choisi par l'utilisateur.

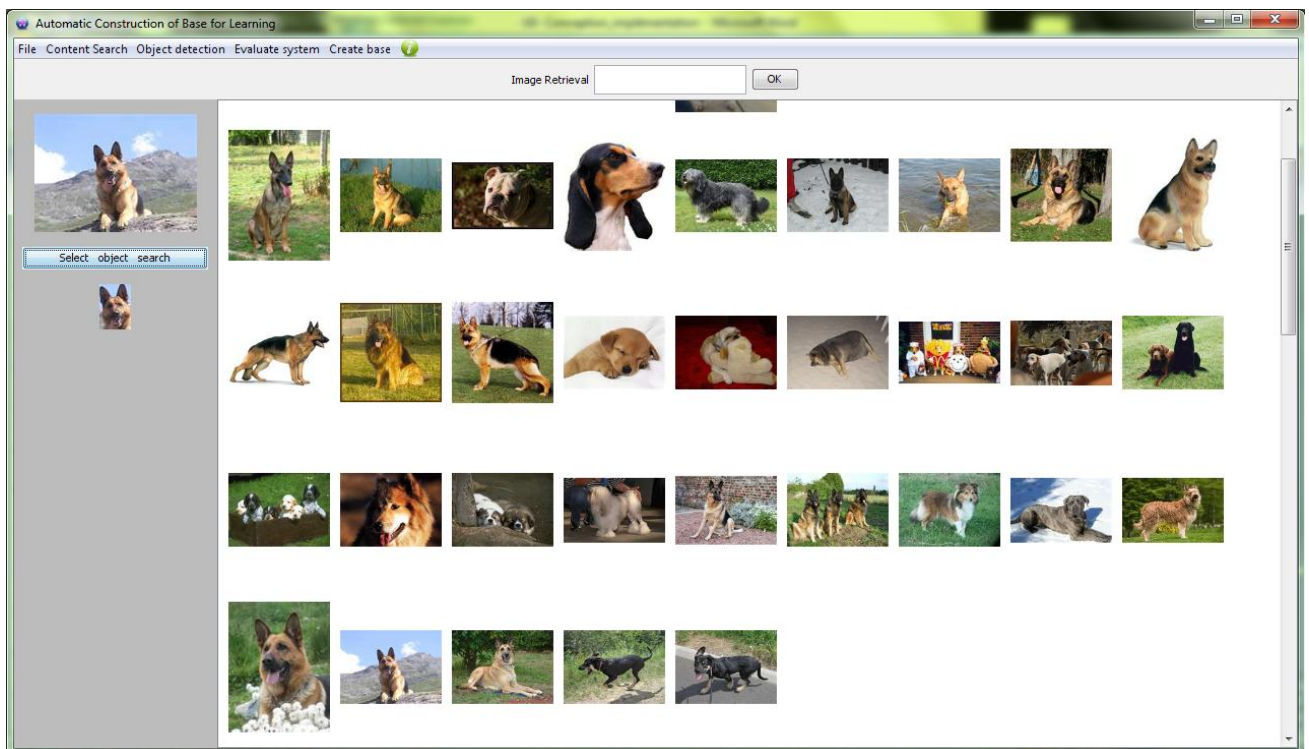


Figure IV.23 : La méthode de la corrélation.

3.4.7. La méthode de la corrélation (variante Sobel)

Dans cette méthode on cherche à laisser que les images qui contiennent l'objet choisi par l'utilisateur en tenant compte qu'on travaille sur l'image contour.

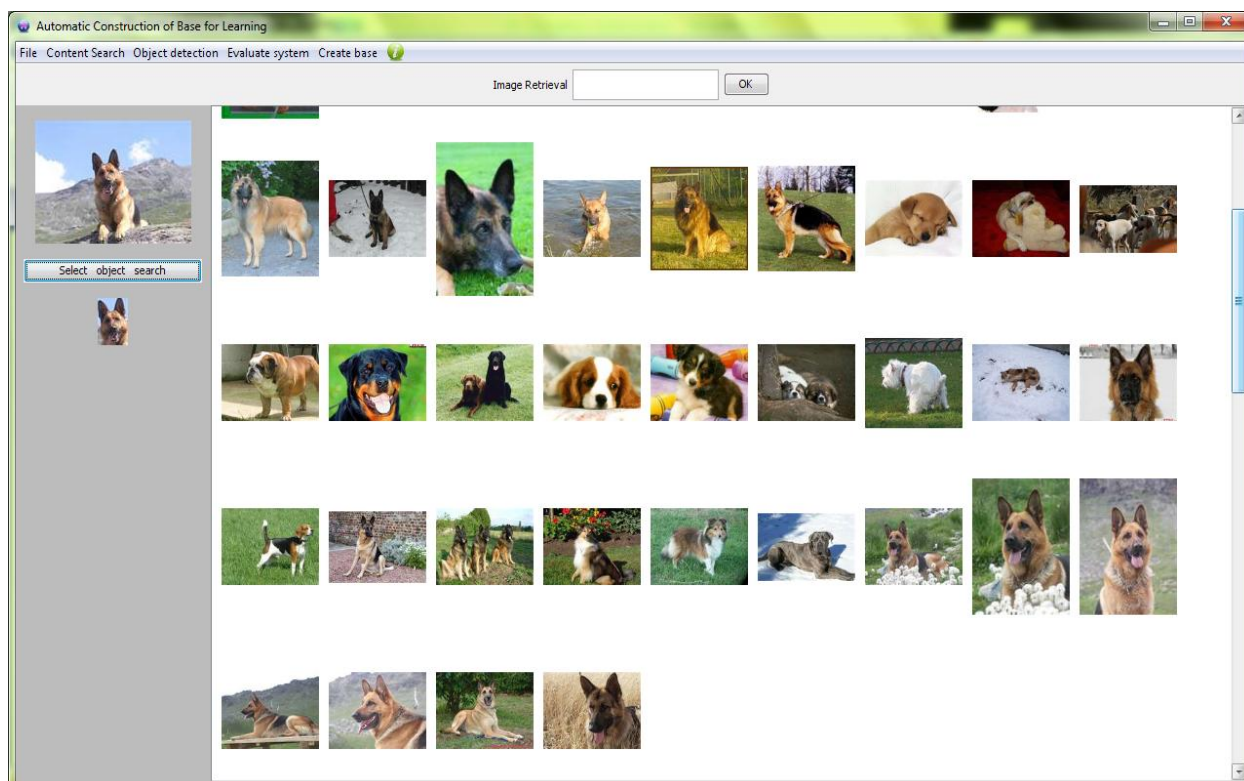


Figure IV.24 : La méthode de la corrélation (variante sobel).

3.4.8. Construction de la base

A chaque fois qu'on applique une de ces dernières méthodes, nous procédons à l'extraction d'attributs caractéristiques pour limiter la quantité d'informations utilisée dans l'apprentissage et améliorer la qualité d'apprentissage. Ici on ne conserve que les attributs caractéristiques discriminants et on élimine ceux redondants et le bruit.

Moyenne	Variance	Enertie	Entropie	Energie	Homogeneité	Classe
Classe chien :						
111.0	70.0	2.0	0.0	1.0	1.0	1
112.0	53.0	2.0	0.0	1.0	1.0	1
133.0	38.0	1.0	0.0	1.0	1.0	1
134.0	50.0	1.0	0.0	1.0	1.0	1
104.0	54.0	1.0	0.0	1.0	1.0	1
115.0	33.0	1.0	0.0	1.0	1.0	1
60.0	76.0	2.0	0.0	1.0	1.0	1
138.0	58.0	1.0	0.0	1.0	1.0	1
144.0	66.0	2.0	0.0	1.0	1.0	1
123.0	61.0	2.0	0.0	1.0	1.0	1
115.0	37.0	1.0	0.0	1.0	1.0	1
123.0	57.0	1.0	0.0	1.0	1.0	1
96.0	41.0	1.0	0.0	1.0	1.0	1
123.0	33.0	1.0	0.0	1.0	1.0	1
105.0	57.0	4.0	0.0	1.0	1.0	1
187.0	64.0	1.0	0.0	1.0	1.0	1
65.0	55.0	1.0	0.0	1.0	1.0	1
138.0	45.0	1.0	0.0	1.0	1.0	1
126.0	52.0	1.0	0.0	1.0	1.0	1
94.0	51.0	2.0	0.0	1.0	1.0	1
126.0	58.0	1.0	0.0	1.0	1.0	1
155.0	71.0	1.0	0.0	1.0	1.0	1

Figure IV.25 : La base d'apprentissage.

4. Etude comparative

4.1. Comparaison des méthodes de recherche d'images par contenu

L'histogramme est une représentation graphique de la distribution des intensités dans une image numérique. Il trace le nombre de pixels pour chaque valeur intensité. Tandis que, la texture est une: "répétition spatiale d'un même motif dans différentes directions de l'espace".

Critère \ Méthode	Histogramme	Texture
Rotation	Invariante	Invariante
Taille d'image	Variante	Invariante
Comparaison	Globale	Globale
Information spatiale	Ne possèdent pas d'informations spatiales sur les positions des couleurs	Possèdent d'informations spatiales sur les positions des couleurs
Changement de luminosité	Sensible	N'est pas sensible

Tableau IV.1 : Comparaison des méthodes de recherche d'images par contenu.

Supposons que l'image requête est de taille 160 * 90, alors les résultats obtenus sont les suivants :

Taille d'image	Distance Manhattan	Distance de texture
Image 160 * 90	0	0
Image 160 * 100	-4800	0
Image 160 * 89	480	0

Tableau IV.2 : Distances en fonction de la taille d'image.

Taille d'image	Distance Manhattan	Distance de texture
L'image originale	0	0
L'image avec position différente (rotation 90°)	0	0

Tableau IV.3 : Distances en fonction de la position d'image.

Nous remarquons que la taille de l'image est un facteur majeur pour calculer la similarité en utilisant l'histogramme de l'image. Elle ne joue aucun rôle dans le cas de la recherche se basant sur la texture. Ces deux méthodes ne sont pas sensibles à la rotation des images.

En appliquant les méthodes de recherche par le contenu sur une des bases téléchargées « chien » qui contient 70 images, nous obtenons les résultats suivants :



Méthodes utilisées pour la recherche	Histogramme	texture	Combinaison
Précision	90,00%	83,33%	94,73%
Rappel	69,23%	96,15%	69,23%
Nombre d'images pertinentes	36/52	50/52	36/52
Temps d'exécutions	703ms	02s 829ms	03s 109ms

Tableau IV.4 : Valeurs de rappel et précision avec tous les descripteurs implémentés.

On remarque que la combinaison des descripteurs améliore le taux de recherche par rapport à l'utilisation d'histogramme ou de texture. Bien que la méthode basée sur la texture donne un nombre important d'images par rapport aux nombre des pertinentes (50/52) et un rappel très performant (96,15% soit la quasi totalité des documents pertinents), elle demeure moins précise que les autres méthodes (83,33% fournira en guise de réponse de nombreux documents erronés en plus de ceux pertinents). Cependant la méthode par combinaison offre

une grande précision avec un taux de rappel égale à celui d’histogramme et inférieur à celui de la texture.

4.2. Comparaison entre les méthodes de détection d’objet et les méthodes de recherche d’image par contenu

Le *Tableau IV.5* représente une comparaison entre les deux méthodes de détection d’objet implémentées :

Critère \ Méthode	Corrélation	Variante
Rotation	Variante	Variante
Translation	Invariante	Invariante
Taille d’image	Invariante	Invariante
Type	Généraliste	Généraliste
Changement de luminosité	Non	Non
Taille d’objet	Sensible	Sensible

Tableau IV.5 : Comparaison des méthodes de détection d’objet.

Le *Tableau IV.6* illustre les valeurs de précision et rappel effectués sur un ensemble d’images récupérées du web « chien » qui contient 70 images.



Image requête :

Méthodes utilisées pour la recherche	Recherche par contenu d’image			Détection d’objet	
	histogramme	Texture	combinaison	Corrélation	Variante
Précision	48,88%	47,69%	49,83%	46,34%	44,89%
Rappel	66,66%	93,93%	63,63%	57,57%	66,66%
Nombre d’images pertinentes/trouvée	22/33	31/33	21/33	19/33	22/33
Temps d’exécutions	703ms	02s 829ms	03s 109ms	37m 11s 894ms	38m 06s 347ms

Tableau IV.6 : Valeurs de rappel et précision avec toutes les méthodes implémentées.

A la lecture du tableau comparative, nous remarquons que les résultats obtenus par contenu d’image sont meilleurs en temps d’exécutions et en qualité d’information.

5. Conclusion

Nous avons présenté un système de construction automatique de bases d'apprentissage « **ACLB : Automatic Construction of Learning Base** », composé de deux parties : la première partie consiste à chercher automatiquement un ensemble d'images du web respectant la requête de l'utilisateur, la deuxième partie joue le rôle d'un filtre sur la base des méthodes implémentées à savoir : les méthodes de recherche par le contenu et les méthodes de détection d'objet. Bien que les deux méthodes présentent des résultats assez comparables. Cependant la deuxième méthode présente l'inconvénient en temps d'exécutions.

*Conclusion générale et
perspectives*

Conclusion générale et perspectives

Dans le cadre de la construction automatique de bases d'apprentissage, beaucoup d'interrogations ont été posées sur les méthodes utilisées pour augmenter la pertinence des résultats de recherche sur le web. A cet effet, trois domaines interviennent:

- ✓ La recherche d'image sur le web qui consiste à récupérer un ensemble d'images qui doivent correspondre à la requête de l'utilisateur.
- ✓ La recherche d'image par le contenu est destinée à réduire l'ensemble des images collectées, et ne garder que celles qui sont vraiment en rapport avec l'image requête.
- ✓ La reconnaissance des formes et détection d'objet utilise des méthodes de détection d'objets pour filtrer les résultats, et ne garder que les images qui contiennent l'objet sélectionné par l'utilisateur dans l'image.

Le but principal de ce travail est la construction automatique de bases d'apprentissage, où nous avons procédé à l'extraction d'attributs caractéristiques pour limiter la quantité d'informations utilisée dans l'apprentissage et améliorer la qualité d'apprentissage en ne gardant que les attributs caractéristiques discriminants et éliminant ceux redondants et le bruit. Ces attributs caractéristiques sont utilisés pour construire la base d'apprentissage. Aussi, nous avons présenté une étude comparative entre les résultats des différentes méthodes implémentées. Cette évaluation facilite à l'utilisateur le choix de la méthode à appliquer pour construire sa base d'apprentissage.

En perspective, nous envisageons utiliser d'autres techniques de recherche d'images par contenu permettant d'augmenter la pertinence des résultats comme les méthodes de recherche par le contenu sémantique de l'image.

Bibliographie

Bibliographie

- [1] Hazem El-Bakry & Kos Mastorakis, « *Fast Information Retrieval from Web Pages* », Computational Intelligence, Man-Machine Systems and Cybernetics, 2008.
- [2] Tollari Sabrina, Detyniecki Marcin & Fakeri-Tabrizi li, « *Utilisation de concepts visuels et de la diversité visuelle pour améliorer la recherche d'images* ».
- [3] Ritendra Datta, Dhiraj Joshi, Jia Li & James z. Wang, « *Image Retrieval: Ideas, Influences, and Trends of the New Age* », ACM Computing Surveys, Vol. 40, No. 2, Avril 2008.
- [4] Christophe Boudry & Clemence Agostini, « *Étude comparative des fonctionnalités des moteurs de recherche d'images sur Internet* », Vol.41, Février 2004.
- [5] M. L. Kherfi, Djemel Ziou & A Bernadi, « *Image Retrieval from the World Wide Web: Issues, Techniques and Systems* », ACM Computing Surveys, Vol.36, No.1, Mars 2004.
- [6] Lina F.Soualmia & Stéfan J.Darmoni, « *Projection de requêtes pour une recherche d'information intelligente sur le Web* ».
- [7] Hichem Bennour, « *Indexation automatique des images : une approche sémantique basée sur l'apprentissage supervisé des régularités* », mémoire de Master, université de Monastir - Tunisie-.
- [8] Dr. Alain Boucher, « *Indexation et recherche d'images par le contenu* », mémoire de Master, 2005.
- [9] H.B.Kekre, Tanuja K. Sarode & Sudeep D. Thepade, « *Image retrieval using color-texture features from DCT on VQ codevectors obtained by kekre's fast codebook generation* », ICGST-GVIP Journal, Vol.9, No.5, Septembre 2009.
- [10] Marine Campedel, « *Indexation des images* », école supérieure des télécommunications - Paris -, Mars 2005.
- [11] Pierre Tirilly, Vincent Claveau & Patrick Gros, « *Annotation d'images sur de grands corpus réels de données* ».
- [12] Hichem Bannour, « *Une approche sémantique basée sur l'apprentissage pour la recherche d'image par contenu* », Conférence en Recherche D'information et Applications, université de Monastir - Tunisie -, 2009.
- [13] Le Thi Lan, « *Interface de visualisation avec retour de pertinence pour la recherche d'images* », 2004.
- [14] Marine campedel, « *Ontologies et représentations symboliques des images* », école supérieure des télécommunications - Paris -, 2006/2007.

- [15] Pr.Seridi Hamid, « *Reconnaissance des formes* », notes de cours, Université 8 Mai 1945, 2010/2011.
- [16] Merabet Fatima, « *Classification Supervisée Par Les Réseaux RBF* », mémoire de Magister, Université 8 Mai 45, 2010.
- [17] Rolf Ingold, « *Reconnaissance des Formes* », notes de cours, université de Fribourg, 2006.
- [18] Damien Poisson & Sami Mahjoub, « *La reconnaissance de forme* », école supérieur de génie informatique – Paris -, 06/06/2010.
- [19] Dr. H.F.Merouani, « *Reconnaissance des formes en imagerie* », notes de cours, université Badji Mokhtar - Annaba -, 2007/2008.
- [20] Marco Treiber, « *An Introduction to object recognition* », Springer, 2010.
- [21] Houaria ABED, Lynda ZAOUÏ & Zakia GUEZZEN, « *Fusion couleur texture dans l'indexation et la recherche des images* », 3èmes Journées Internationales sur l'Informatique Graphique, université des sciences et de la technologie d'Oran - Mohamed Boudiaf -, 2007.
- [22] Eugene K. yen & Roger G. Johnston, « *The ineffectiveness of the correaltion coefficient for image comparisons* », Los Alamos National Laboratory.
- [23] Antoine Mirecourt, « *Le développeur Java 2* », édition 1999.
- [24] Loafi Wafa & Bouchreb Lotfi, « *Conception et Réalisation d'un Système d'Aide pour l'Interprétation des Mammographies SIAM* », mémoire d'ingénieur, université 8 Mai 45, Juin 2008.
- [25] Toualbia Ilyes & Djellab Issam, « *Vers une approche de segmentation incrémentale pour l'extraction des caractéristiques du visage* », mémoire d'ingénieur, université 8 Mai 45, Juin 2005.

Webographie :

- [web 1] http://fr.wikipedia.org/wiki/Inf%C3%A9rence_bay%C3%A9sienne, 12/05/11.
- [web 2] http://fr.wikipedia.org/wiki/Classifieur_lin%C3%A9aire, 12/05/11.
- [web 3] http://fr.wikipedia.org/wiki/R%C3%A9seau_de_neurones, 12/05/11.
- [web 4] http://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support,12/05/11.
- [web 5] <http://docs.gimp.org/fr/plugin-sobel.html>, 13/05/11.
- [web 6] http://fr.wikipedia.org/wiki/Algorithme_de_Sobel, 13/05/11.
- [web 7] <http://www.commentcamarche.net/contents/video/traitimg.php3>, 13/05/11.

Annexe

Annexe

Formule 1 : calcul du rappel

Le ratio entre le nombre de documents pertinents retrouvés et le nombre total de documents pertinents dans la base.

$$rappel_q = \frac{R_q \cap P_q}{P_q}$$

Avec :

R_q : est l'ensemble des documents retrouvés par le système pour la requête q

P_q : est l'ensemble des documents pertinents de la collection pour cette requête.

Formule 2 : calcul de la précision

Le nombre de documents pertinents retrouvés rapporté au nombre total de documents retrouvés.

$$precision_q = \frac{R_q \cap P_q}{R_q}$$

Avec :

R_q : est l'ensemble des documents retrouvés par le système pour la requête q

P_q : est l'ensemble des documents pertinents de la collection pour cette requête.

Formule 3 : calcul de la moyenne

$$Moy_I = \frac{1}{N * M} * \sum_{i=0}^{imax} n(i) * i$$

Avec :

Moy_I : Moyenne de l'image I

N, M : largeur et hauteur de l'image I

$imax$: intensité maximale

$n(i)$: nombre d'occurrence de l'intensité i

Formule 4 : calcul de la variance

$$Var_I = \sqrt{\frac{1}{N * M} * \sum_{i=0}^{imax} n(i) * (i - Moy_I)^2}$$

Avec :

Var_I : variance de l'image I

N, M : largeur et hauteur de l'image I

i_{max} : intensité maximale

$n(i)$: nombre d'occurrence de l'intensité i

Formule 5 : Calcul de l'inertie (contraste)

$$\text{Inertie} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} ((i-j)^2 P_{ij}(d, \theta))$$

Avec :

P_{ij} : La matrice de cooccurrence.

d : la distance.

θ : La direction.

Formule 6 : Calcul de l'énergie

$$\text{Energie} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (P_{ij}(d, \theta)^2)$$

Avec :

P_{ij} : La matrice de cooccurrence.

d : la distance.

θ : La direction.

Formule 7 : Calcul de l'entropie

$$\text{Entropie} = - \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\log P_{ij}(d, \theta) P_{ij}(d, \theta))$$

Avec :

P_{ij} : La matrice de cooccurrence.

d : la distance.

θ : La direction.

Formule 8 : Calcul de l'homogénéité (moment inverse de différence)

$$\text{Homogénéité} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \left(\frac{P_{ij}(d, \theta)}{1 + |i-j|^2} \right), \text{ avec } i \neq j$$

Avec :

P_{ij} : La matrice de cooccurrence.

d : la distance.

θ : La direction.