

**UNIVERSITE KASDI MERBAH OUARGLA**  
**Faculté des Sciences et de La Technologie et Sciences de La Matière**  
**Département de Mathématique Et d'Informatique**



**Mémoire**  
**MASTER ACADEMIQUE / PROFESSIONNEL**

Domaine : Mathématique Et Informatique  
Filière : Informatique Académique  
Spécialité. Informatique Fondamental

Présenté par : Taibaoui Mohamed  
Debbar Djafar

**Thème**

**La découverte des concepts sémantiques cachés avec  
plusieurs niveaux d'abstraction pour la recherche  
d'images**

Soutenu publiquement  
En : juin 2013

Devant le jury :

Mr. MEFLEH Mohammed Salim	MA(A)	Président	UKM Ouargla
Melle KORICHI Meriem	MA(B)	Examineur	UKM Ouargla
Mme DEBBAGH Farah	MA(A)	Encadreur	UKM Ouargla

**Année Universitaire : 2012 /2013**

# *Remerciements*

*Nous remercions DIEU le tout puissant, maître des cieux et de la terre, qui nous a éclairé le chemin et permis de mener à bien ce travail.*

*Tout d'abord nous tenons surtout à adresser nos plus vifs remerciements à Mme Debbagh Farah, qui nous a permis de réaliser ce travail sous sa direction. Nous oublierons jamais ses conseils judicieux.*

*Un merci pour tous mes enseignants surtout monsieur Meflah mohammed Salim qui nous a permis de travailler dans un environnement serein et paisible, sans oublier monsieur Euschi salah et khaldi bilel Qui nous ont aidé beaucoup*

*Aux membres du jury pour avoir accepté d'évaluer notre travail.*

*Un grand merci à toutes les personnes qui nous ont soutenues de près ou de loin au cours de la réalisation de ce modeste travail.*

# *Dédicaces*

*Je dédie ce travail*

*A ceux qui sont dans mon cœur, qui ont veillés pour notre confort et sacrifié  
beaucoup pour notre réussite, Ma chère mère, mes chères grand-mères  
jamila et fatima (que dieu me les garde)*

*A celui qui m'a toujours appris comment réfléchir avant d'agir, à celui qui m'a  
soutenu tout au long de ma vie scolaire, à celui qui n'a jamais épargné un effort  
pour mon bien, Mon cher père*

*(Que dieu me le garde)*

*A mes chers frères, mes chères sœurs, toutes ma famille, mes chères amis*

*A tous les membres de la promotion master informatique de l'université Kasdi  
Merbeh OUARGLA.*

*Et à tous ceux qui me connaissent*

***A tous qui m'aiment et je l'aime R***

***Mohamed***

# *DEDICACES*

*A ma mère et mon père que DIEU les gardent.*

*A mes sœurs, mes frères.*

*A tous ceux qui aiment DJAAFAR.*

*A tous ceux qui sont proches de mon cœur.*

*A tous mes amis.*

*DJAAFAR*

## Résumé

Les images numériques constituent une source d'informations très expressive, et qui joue un rôle très important dans des nombreuses activités humaines. Par conséquent, Avoir des volumes croissants des bases d'images est un résultat naturel. A cet effet, et pour assurer une exploitation efficace des images par des machines, leur description est une tâche qui s'impose. Sinon, ces images peuvent demeurer pratiquement inexploitable et impossible de les retrouver et utiliser efficacement.

L'annotation des images est un domaine de recherche très actif. Elle consiste à désambiguïser et interpréter les images pour des traitements automatiques variés tels que : la recherche et la classification. Les travaux autour de ça, on s'intéressés par deux problématiques : **comment faire cette description ? Et par quoi faire ?**

Dans le cadre de notre travail nous somme intéressé, pour répondre à la première problématique, par la proposition et l'implémentation d'une méthode d'annotation semi-automatique. Ainsi que, pour répondre à la deuxième problématique, nous avons opté pour l'annotation sémantique à base d'ontologie pour profiter de la richesse sémantique quelle présente. Nous avons choisis la nature comme domaine pour notre ontologie orienté terminologie. Notre travail d'annotation semi-automatique consiste à :

- Segmenter chaque image de la base en des régions. Afin de pouvoir, par la suite, connaître ses constituants.
- Faire une annotation manuelle d'un échantillon de régions par des concepts. De telle sorte que nous couvrions tous les concepts sémantiques.

- Ensuite, procéder l'annotation automatique pour le reste des régions par propagation de concepts sémantiques en se basant sur des mesures de similarité visuelles entre régions. Et enfin, nous récupérerons les concepts de chaque image, en fonction de ses régions.

**Mots clés : image numérique, segmentation, région, interprétation, ontologie, annotation sémantique.**

## ملخص

الصور الرقمية هي مصدر للمعلومات معيرة جدا، وتلعب دورا هاما في كثير من الأنشطة البشرية. وبناء على ذلك، وجود كميات متزايدة من قواعد الصور هو نتيجة طبيعية. وتحقيقا لهذه الغاية، وضمان الاستخدام الفعال لهذه الصور من قبل الأجهزة، وصفهم هو المهمة المرجوة. خلاف ذلك، قد تبقى هذه الصور غير صالحة للاستعمال من الناحية العملية ومن المستحيل العثور عليها واستخدامها بشكل فعال.

شرح الصور هو مجال بحث نشط جدا. يعتمد على توضيح وتفسير الصور من أجل معالجات آلية مختلفة مثل البحث والتصنيف. هذه الأعمال تهتم بإشكاليتين: كيفية القيام بهذا الوصف؟ وبمذا يتم؟

في عملنا نحن مهتمون بالإجابة على المشكلة الأولى، باقتراح وتنفيذ طريقة الشرح شبه التلقائي. وللإجابة عن المشكلة الثانية، اخترنا الشرح الدلالي القائم على الأنطولوجيا من أجل الاستفادة من ثراءها الدلالي. اخترنا الطبيعة ك مجال للأنطولوجيا. عمل الشرح شبه التلقائي الذي قمنا به يعتمد على :

- تقطيع كل صورة من قاعدة الصور الى مناطق لمعرفة مكوناتها لاحقا.
- وضع شرح يدوي لعينة من المناطق باستعمال المفاهيم الدلالية. بحيث تتم تغطية جميع المفاهيم الدلالية.
- ثم تجري الشرح التلقائي للمناطق المتبقية بواسطة نشر المفاهيم الدلالية بالاعتماد على مقاييس التشابه البصري بين المناطق. وأخيرا، نحصل على مفاهيم كل صورة، اعتمادا على مناطقها.

**كلمات مفتاحية : الصور الرقمية، تقسيم الصور الرقمية، المنطقة، والتفسير، الأنطولوجيا ، والشرح الدلالي.**

## Abstract

Digital images are a very expressive information source, and which plays a very important role in many human activities. As a result, growing volumes of the bases of images is a natural result. To this end, and to ensure effective use of images by machines, their description is a required task. Otherwise, these images can remain virtually useless and impossible to find them and use effectively.

The annotation of images is a very active area of research. It is to disambiguate and interpret the images for automatic and varied treatments such as: research and classification. Work around it, is interested in two issues: how to make this description? And with what to do?

In our work we are interested, to address the first problem, by proposal and implementation of a semi-automatic annotation method. So, to answer the second question, we have opted for ontology-based semantic annotation to take advantage of the semantic richness witch present. We selected the nature as a domain for our ontology-oriented terminology. Our semi-automatic annotation work is to:

- Segment each image of the base areas. In order to, subsequently, know its constituents.
- Make a manual annotation of a few regions by concepts. So we will cover all the semantic concepts.
- Then proceed the automatic annotation for the rest of the regions by spreading semantic concepts based on Visual similarity measurements between regions. And finally, we recover the concepts of each image, in function of its regions.

**Key words: digital image, segmentation, region, interpretation, ontology, semantic annotation.**

# SOMMAIRE

<i>CHAPITRE I : Introduction générale</i> .....	1
1. Introduction.....	1
2. Motivations.....	1
3. Contributions.....	2
4. Organisation du mémoire.....	3
<i>CHAPITRE II : La segmentation d'images</i> .....	5
1. Introduction.....	6
2. Image et caractéristiques.....	6
2.1. Définition de l'image numérique.....	6
2.2. Caractéristiques de l'image numérique.....	7
2.2.1. Un pixel.....	7
2.2.2. Une région.....	8
2.2.3. La dimension.....	8
2.2.4. La résolution.....	8
2.2.5. La Couleur.....	9
2.2.6. La texture.....	11
2.2.6.1. Définition.....	11
2.2.6.2. La matrice de co-occurrence.....	11
2.2.7. La forme.....	12
3. La segmentation.....	13
3.1. Définition de la segmentation d'images.....	13
3.2. Les objectifs de la segmentation.....	14
3.3. Les vues de la segmentation.....	14
3.3.1. Segmentation selon les connaissances a priori.....	14
3.3.1.1 Segmentation supervisée.....	14
3.3.1.2 Segmentation non supervisée.....	15
3.3.2. Segmentation selon l'espace d'image.....	15
3.3.2.1. Segmentation par approche contour.....	15
3.3.2.2. Segmentation par approche région.....	16
4. Conclusion.....	17
<i>CHAPITRE III : Annotation et Ontologie</i> .....	18
1. Introduction.....	19
2. Annotations et Métadonnées.....	19
2.1. Définitions.....	19
2.1.1. Métadonnée.....	19
2.1.2. Annotation.....	20
2.1.3. Annotation sémantique.....	20
2.2. Méthodes d'annotation.....	21
2.2.1. L'annotation manuelle.....	21
2.2.2. L'annotation semi-automatique.....	22

2.2.3.	L'annotation automatique .....	22
2.3.	Natures de métadonnées .....	23
2.3.1.	Métadonnée indépendante du contenu .....	23
2.3.2.	Métadonnées dépendantes de contenu .....	23
2.3.2.1.	Métadonnées dépendantes de contenu symbolique .....	23
2.3.2.2.	Métadonnées descriptives de contenu sémantique .....	23
3.	Les ontologies .....	24
3.1.	Définition d'une ontologie.....	25
3.2.	Les motivations du développement des ontologies .....	25
3.3.	Les composantes des ontologies .....	26
3.3.1.	Les concepts .....	26
3.3.2.	Les classes .....	26
3.3.3.	Les attributs .....	27
3.3.4.	Les relations .....	27
3.3.5.	Les axiomes .....	27
3.4.	Le cycle de vie des ontologies .....	27
3.5.	L'ingénierie ontologique .....	28
3.5.1.	Spécification des besoins .....	29
3.5.2.	Conceptualisation .....	29
3.5.3.	Formalisation.....	30
3.5.4.	Implémentation.....	30
3.6.	Les langages des ontologies .....	30
3.7.	Les environnements et des outils de développement d'ontologies .....	32
3.7.1.	Ontolingua.....	32
3.7.2.	Ontosaurus.....	32
3.7.3.	ODE.....	32
3.7.4.	Protégé.....	32
4.	Conclusion .....	34
<i>CHAPITRE IV : Conception</i> .....		35
1.	Introduction.....	36
2.	Conception de notre ontologie « Onto_Nature» .....	36
	Phase 1: Spécification des besoins .....	36
	Phase 2 : Conceptualisation.....	37
3.	Conception de notre système d'annotation sémantique semi-automatique.....	40
3.1.	Le choix d'une base d'images .....	41
3.2.	La segmentation de la base d'images .....	41
3.2.1.	Le filtrage des images.....	43
3.2.2.	La segmentation des images filtrées .....	43
3.3.	L'extraction des caractéristiques des régions .....	44
3.3.1.	Le calcul de la couleur .....	44
3.3.2.	Le calcul de la texture.....	45
3.4.	Module d'annotation sémantique semi-automatique .....	46
3.4.1.	Le sous module d'annotation manuelle.....	48
3.4.2.	Le sous module d'annotation automatique .....	48
3.4.2.1.	Comparaison de la couleur .....	48

3.4.2.2.	Comparaison de la texture .....	49
3.4.2.3.	Comparaison de la combinaison couleur et texture.....	49
3.4.3.	Association des concepts aux images .....	50
4.	Conclusion .....	51
<i>CHAPITRE V : Réalisation</i> .....		52
1.	Introduction.....	53
2.	Implémentation de notre ontologie « <i>onto_nature</i> » orienté terminologie .....	53
2.1.	Choix de l'éditeur d'ontologie.....	53
2.2.	Implémentation de notre ontologie .....	54
2.2.1.	Création des classes .....	54
2.2.2.	Création des relations .....	54
2.2.3.	Visualisation de l'ontologie.....	55
3.	L'implémentation de notre système d'annotation sémantique semi-automatique des images....	56
3.1.	Choix des outils .....	56
3.1.1.	Visual Studio .....	56
3.1.2.	C Sharp (c#).....	57
3.1.3.	DevExpress.....	57
3.1.4.	Adobe Photoshop.....	57
3.2.	Les étapes d'implémentation.....	57
3.2.1.	Le choix d'une base d'images .....	57
3.2.2.	La segmentation de la base d'images .....	57
3.2.2.1.	L'implémentation de filtrage d'images.....	57
3.2.2.2.	Implémentation de la segmentation des images filtrées .....	58
3.2.3.	L'interface principale de notre application .....	60
3.2.4.	Le calcul des caractéristiques visuelles des régions .....	61
3.2.4.1.	Calcul de l'histogramme RGB de chaque région .....	61
3.2.4.2.	Le Calcul de la matrice de cooccurrence pour chaque région .....	62
3.2.5.	Implémentation de l'annotation sémantique semi-automatique .....	65
3.2.5.1.	La création d'un dictionnaire .....	65
3.2.5.2.	L'annotation manuelle d'un échantillon de régions .....	65
3.2.5.3.	L'annotation automatique des régions restantes.....	67
3.2.5.4.	Association des concepts aux images .....	74
4.	Conclusion .....	75
<i>Conclusion générale</i> .....		76
<i>BIBLIOGRAPHIE</i> .....		78
<i>WEBOGRAHIE</i> .....		79



# LISTE DE FIGURE

Figure - 1 : La lettre A.....	7
Figure - 2 : Dimension d'une image.....	8
Figure - 3 : Résolution d'une image.....	8
Figure - 4 : Les trois primaires: Rouge, Vert, et Bleu.....	9
Figure - 5 : TSL (teinte, saturation, luminosité) sous forme cône.....	10
Figure - 6 : Exemple d'un histogramme.....	10
Figure - 7 : Des exemples de textures.....	11
Figure - 8 : Les étapes de traitement d'images.....	13
Figure - 9 : les vues de segmentation d'images.....	14
Figure - 10 : Les approches de segmentation selon l'espace d'image.....	15
Figure - 11 : Segmentation par la méthode de contour actif.....	16
Figure - 12 : Une segmentation obtenue grâce à LPE.....	16
Figure - 13 : La segmentation de type split & merge.....	16
Figure - 14 : Cycle de vie d'ontologie.....	28
Figure - 15 : Langage traditionnel d'ontologies.....	30
Figure - 16 : Langages d'ontologies.....	31
Figure - 17 : l'interface de Protégé.....	33
Figure - 18 : Hiérarchie de notre ontologie « Onto_Nature ».....	39
Figure - 19 : Architecture générale de notre système.....	41
Figure - 20 : Architecture de module de segmentation.....	43
Figure - 21 : Les modules des calculs des caractéristiques des régions.....	44
Figure - 22 : calculer de l'histogramme d'une région.....	44
Figure - 23 : La structure de fichier de caractéristique.....	46
Figure - 24 : Architecture de l'annotation sémantique.....	47
Figure - 25 : La structure de fichier d'annotation de région.....	50
Figure - 26 : La structure de fichier d'annotation des images.....	50
Figure - 27 : Editeur d'ontologie Protégé.....	53
Figure - 28 : Description des Classes d'onto_nature.....	54
Figure - 29 : Création des relations.....	55
Figure - 30 : Visualisation de l'ontologie par Protégé.....	55
Figure - 31 : le code OWL de notre ontologie « Onto_Nature ».....	56
Figure - 32 :L'interface de filtrage.....	58
Figure - 33 : Sélection des régions via l'outil « Quick Selection Tool ».....	58
Figure - 34 : Une région avec un arrière-plan blanc.....	59
Figure - 35 : Le résultat de segmentation, la base de régions.....	59
Figure - 36 : Interface pour lancer l'application.....	60
Figure - 37 : l'interface principale.....	60
Figure - 38 : l'interface de calculer les caractéristiques des régions.....	64
Figure - 39 : un exemple de fichier des caractéristiques.....	65
Figure - 40 : dictionnaire vide.....	65
Figure - 41 : Un échantillon de régions dans le concept ciel du dictionnaire.....	66
Figure - 42 : Un échantillon de régions dans le concept palmier du dictionnaire.....	66
Figure - 43 : Dictionnaire avant la propagation des concepts.....	67
Figure - 44: Résultat Après mesure de similarité visuelle de la couleur.....	69

Figure - 45 : Le contenu du concept arbre avant. ....	71
Figure - 46 : Résultat Après mesure de similarité visuelle de la texture .....	71
Figure - 47: Le contenu du concept arbre avant .....	72
Figure - 48: Résultat Après mesure de similarité visuelle par la combinaison .....	73
Figure - 49: Le fichier d'annotation résultat.....	74
Figure - 50 : la base d'images avant et après l'annotation semi-automatique.....	75

## **List des Tableaux**

Tableau 1 : Classes et hiérarchie des classes de l'ontologie .....	31
Tableau2 : Sélection des caractéristiques et des poids pour chaque concept .....	66

# **CHAPITRE I : Introduction générale**

- **Introduction**
- **Motivations**
- **Contributions**
- **Organisation du mémoire**

# CHAPITRE I : Introduction générale

---

## 1. Introduction

Avec le développement de l'Internet et la disponibilité des captures d'images via des dispositifs tels qu'une caméra numérique et le scanner, la taille des bases d'images numériques est en augmentation très rapide dans des domaines variés tel que: la médecine, l'archives (patrimoine culturel, musées), les agences photographiques, l'éducation, l'image satellitaire et aérienne,...etc.

Par conséquence, et pour assurer une exploitation efficace de ces bases d'images par des machines, la description de telles bases est une tâche qui s'impose. Sinon, ces images peuvent demeurer pratiquement inexploitable et impossible de les retrouver et utiliser efficacement.

L'annotation des images est un domaine de recherche très actif. Elle consiste à désambiguïser et interpréter les images d'une façon qui assure leur exploitation adéquate par la machine pour des buts variées tels que : la recherche et la classification.

## 2. Motivations

La description des images est une tâche indispensable pour une exploitation efficace par la machine. Pour cela, le besoin à des techniques d'indexation et d'interprétation des images est vite apparu comme fondamental est incontournable. Les travaux autour de ça, ont s'intéressés par deux problématiques : **comment faire cette description ? Et par quoi faire ?**

La réponse à la première problématique, c'est qu'ils existent trois méthodes d'annotation d'images : manuelle, semi-automatique et automatique.

L'annotation manuelle est entièrement effectuée par un annotateur humain chargé d'attribuer à chaque image des annotations selon son point de vue. Elle est acceptable dans le cas d'une petite collection d'images. Cependant, pour un cas réel où les bases d'images sont en augmentation contenue, l'annotation manuelle prene beaucoup de temps ou devenue presque impossible. Par conséquent il faut penser à une autre solution.

Pour L'annotation automatique, elle est entièrement réalisée par une machine et vise à alléger la charge de l'utilisateur. Cependant, il n'existe pas à l'heure actuelle des méthodes automatiques proprement dites.

Afin de faire un compromis entre ces deux méthodes, leur combinaison est devenue nécessaire. C'est ce qui est connu sous le nom "l'annotation semi-automatique".

# CHAPITRE I : Introduction générale

---

La réponse à la deuxième question, c'est qu'ils existent des méthodes de description visuelle et des méthodes de description textuelle.

Les méthodes de description à base de contenu visuel (méthodes de description de contenu symbolique) sont très utilisées dans la littérature. Elles se basent sur le contenu visuel des images telle que la couleur et la texture, et construisent des descripteurs dits visuels. Par la suite, ce sont ces descripteurs qui sont exploités pour les différents traitements tels que la recherche d'images.

Ces méthodes ont prouvé leur réussite, pour les traitements qui s'intéressent par l'aspect visuel seulement. Cependant, elles restent insuffisantes, dans d'autres traitements, car elles ne donnent aucune interprétation des images. Pour cela, le recours à d'autres techniques était indispensable. Ce sont les méthodes de descriptions à base de texte.

La méthode de description textuelle connue, et utilisée par les moteurs de recherches, est l'annotation à base de mots clés. Chaque ressource est annotée par des mots-clés qui lui sont associés. Cette méthode a posé rapidement le problème majeur de bruit et silence.

Pour cela, une nouvelle technique est apparue pour améliorer le résultat d'un traitement sur les images et surtout la recherche. C'est l'annotation sémantique à base d'ontologie. Ces techniques consistent à interpréter une image ou donner sa signification à partir d'une ontologie. Elles sont semblables à la description à base de mots-clés. Cependant la différence c'est que, d'une part, elles utilisent des concepts dans une ontologie. Et d'autre part qu'il ya des relations sémantiques entre les concepts, ce qui permet de enrichir les annotations.

## 3. Contributions

Dans le cadre de notre travail nous sommes intéressés, pour répondre à la première problématique, par la proposition et l'implémentation d'une méthode d'annotation semi-automatique. Ainsi que, pour répondre à la deuxième problématique, nous avons opté pour l'annotation sémantique à base d'ontologie pour profiter de la richesse sémantique qu'elle présente.

Nous présentons dans ce qui suit nos principales contributions :

1. D'abord, nous avons étudié les différentes méthodes d'annotations, ainsi que les différentes formes que peut avoir une annotation. Sur la base de cette étude, nous avons choisi l'annotation semi-automatique à base d'ontologie.

## CHAPITRE I : Introduction générale

---

2. Par conséquent, nous allons construire une ontologie de domaine *Onto-Nature*. C'est une ontologie orienté terminologie sur le domaine de la nature. Elle sera exploitée dans la phase d'annotation.

3. Notre annotation semi-automatique consiste à :

- Segmenter chaque image de la base en des régions. Afin de pouvoir, par la suite, connaître ses constituants.
- Faire une annotation manuelle d'un échantillon de régions. De telle sorte que nous devrions couvrir tous les concepts sémantiques.
- Ensuite, lancer l'annotation automatique pour le reste des régions par propagation de concepts sémantiques en se basant sur des mesures de similarité visuelles entre régions.
- Et enfin, récupérer les concepts de chaque image.

### 4. Organisation du mémoire

Notre mémoire est organisée en cinq chapitres incluant cette Introduction et se termine par une Conclusion générale :

**Le chapitre 2 :** Nous dressons un état de l'art du domaine de la segmentation d'images. Nous commençons dans une première section, par présentation de l'image numérique et ses principales caractéristiques. Ensuite dans une deuxième section, nous commençons par une définition de la segmentation d'images, ses objectifs. Ensuite, les vues de la segmentation d'images.

**Le chapitre3:** Est consacré pour permettre au lecteur de comprendre le sujet d'annotation et l'utilisation des ontologies dans ce sens. Pour cela, nous allons consacrer une première section afin de présenter des définitions indispensables pour distinguer entre les termes métadonnée, annotation et annotation sémantique. Ensuite, nous allons montrer les différents types et les différentes méthodes d'annotations. Dans la deuxième section, nous faisons un tour d'horizon sur les ontologies, comme modèle de représentation de connaissances.

**Le chapitre 4:** C'est dans ce chapitre que nous dériverons la démarche conceptuelle de notre application. Pour cela, nous allons commencer par la conception de notre ontologie *Onto-Nature* sur le domaine de la nature. Ensuite, Nous présentons la conception détaillée de notre système d'annotation sémantique semi-automatique des images, qui va exploiter notre ontologie pour interpréter les images.

## CHAPITRE I : Introduction générale

---

**Le chapitre 5:** Nous consacrons ce chapitre pour décrire la réalisation de notre application. Nous commençons tout d'abord par une motivation des choix des outils pour l'implémentation de notre ontologie « *Onto\_Nature* » et pour l'implémentation de notre application. Par la suite, nous présentons la réalisation de notre ontologie. Par la suite, nous détaillerons le processus d'implémentation de notre système d'annotation sémantique semi-automatique des images ainsi que des captures d'écrans des différentes interfaces et résultats obtenus.

En conclusion nous présenterons les principaux points de ce travail et quelques perspectives qui peuvent se décollées.

## **CHAPITRE II : La segmentation d'images**

- **Introduction**
- **Image et caractéristiques**
- **La segmentation**
- **Conclusion**



### 1. Introduction

L'image est une représentation d'une personne ou d'un objet par la peinture, la sculpture, le dessin, la photographie ou le film, etc. Elle constitue un moyen de communication universel dont la richesse du contenu permet aux êtres humains de tout âge et de toute culture de se comprendre. Chacun peut analyser l'image à sa manière, pour en extraire des informations précises.

L'analyse (ou le traitement) d'images est une discipline qui étudie les images numériques et leurs transformations, dans le but d'améliorer leurs qualités ou d'en extraire de l'information. Parmi les traitements qui existent, il ya la segmentation.

La segmentation d'images est une étape fondamentale et importante dans beaucoup d'applications de vision par ordinateur. C'est une étape primordiale pour l'interprétation d'images. Le but de segmentation est l'extraction d'attributs caractérisant ces entités. Les attributs étudiés correspondent à des points d'intérêt ou à des zones caractéristiques de l'image.

Le but de ce chapitre est de donner l'essentiel sur l'image numérique et la segmentation d'images. Il est organisé en deux sections comme suit :

Dans une première section, sur l'image numérique, nous donnons une définition de l'image numérique et ses principales caractéristiques qui nous intéressent dans notre travail.

Dans la deuxième section, nous allons commencer par une définition de la segmentation d'images, les objectifs de la segmentation. Ensuite, les vues de la segmentation d'images. Enfin, nous terminons ce chapitre par une conclusion.

### 2. Image et caractéristiques

Avant d'aborder le sujet de la segmentation, il est indispensable de présenter des notions de base sur l'image numérique.

#### 2.1. Définition de l'image numérique

La numérisation d'une image est le processus qui permet de passer d'une image physique (image optique par exemple) qui est caractérisée par l'aspect continu du signal qu'elle représente (une infinité de valeur dans l'intensité lumineuse par exemple), à l'état d'image numérique qui est caractérisée par l'aspect discret (l'intensité lumineuse ne peut prendre que des valeurs quantifiées en un nombre fini de points distincts). [Site 1]

## CHAPITRE II : La segmentation d'images

---

Par conséquent, une image numérique est une image (dessin, icône, photographie, ...) créée, traitée, stockée sous forme binaire (suite de 0 et de 1). C'est cette forme numérique qui permet une exploitation ultérieure par des outils logiciels. **[Site 1]**

Lorsqu'on agrandit une image numérique, on voit que celle-ci est composée d'un ensemble de points, appelés « pixels », est désigné par ses coordonnées entières. **[Guediri, K, 12]**

### 2.2. Caractéristiques de l'image numérique

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

#### 2.2.1. Un pixel

Contraction de l'expression anglaise " Picture Elements ": éléments d'image, le pixel est le plus petit point de l'image, c'est une valeur numérique représentative des intensités lumineuses.

Si le bit est la plus petite unité d'information que peut traiter un ordinateur, le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels sur l'image. La lettre A, par exemple, peut être affichée comme un groupe de pixels dans la figure ci-dessous. **[Site 2]**



**Figure - 1 : La lettre A**

La quantité d'information que véhicule chaque pixel donne des nuances entre images monochromes et images couleurs. Dans le cas d'une image monochrome, chaque pixel est codé sur un octet, et la taille mémoire nécessaire pour afficher une telle image est directement liée à la taille de l'image. Cependant, Dans une image couleur (R.V.B.), un pixel peut être représenté sur trois octets : un octet pour chacune des couleurs : rouge (R), vert (V) et bleu (B). **[Site 2]**

### 2.2.2. Une région

La notion de région dans le traitement d'images, est de regrouper des zones possédant les mêmes caractéristiques. C'est-à-dire que si plusieurs pixels adjacents s'avèrent être de couleur identique alors la zone qu'ils forment est une région. [Guediri. K, 12]

Une région peut être définie aussi comme un ensemble de pixels connexes ayant des propriétés communes qui les différencient des pixels des régions voisines. [Do.M.C, 07]

### 2.2.3. La dimension

L'image numérique se présente sous forme d'une matrice dont les éléments sont des pixels. La dimension (ou la taille) d'une image est le nombre de lignes de cette matrice multiplié par le nombre de colonnes, et qui nous donne le nombre total de pixels dans cette image. [Site 2]



Figure - 2 : Dimension d'une image

### 2.2.4. La résolution

La résolution d'une image correspond au niveau de détail qui va être représenté sur cette image. C'est le nombre de pixels par unité de longueur dans l'image à numériser. Elle est en dpi (dots per inch) ou en ppp (points par pouce).

Plus le nombre de pixels est élevé par unité de longueur de l'image à numériser, plus la quantité d'information qui décrit l'image est importante et plus la résolution est élevée. [Site 2]



Figure - 3 : Résolution d'une image

### 2.2.5. La Couleur

La couleur est l'un des plus reconnaissables éléments du contenu visuel d'une image, c'est une caractéristique très utilisée dans la recherche d'images. Il existe plusieurs descripteurs de couleur tel que: les espaces de couleur est l'histogramme. [Ben Cheikh.N et al, 11]

- **Espace de couleur** : une image est composée de pixels. Chaque pixel d'une image peut être représenté comme un point dans un espace de couleur à 3 dimensions (généralement). Les plus utilisés sont: RVB (Rouge Vert Bleu), TSL (Teinte, Saturation, Luminosité)...etc [Ben Cheikh.N et al, 11]

- **RGB**: C'est le modèle de couleur le plus utilisé pour la représentation de la couleur. Il est composé des trois couleurs primaires : rouge, vert et bleu. Ce modèle est aussi le plus utilisé pour la reproduction de la couleur sur les dispositifs d'affichage tels que la télévision et les écrans des ordinateurs. [Ben Cheikh.N et al, 11]

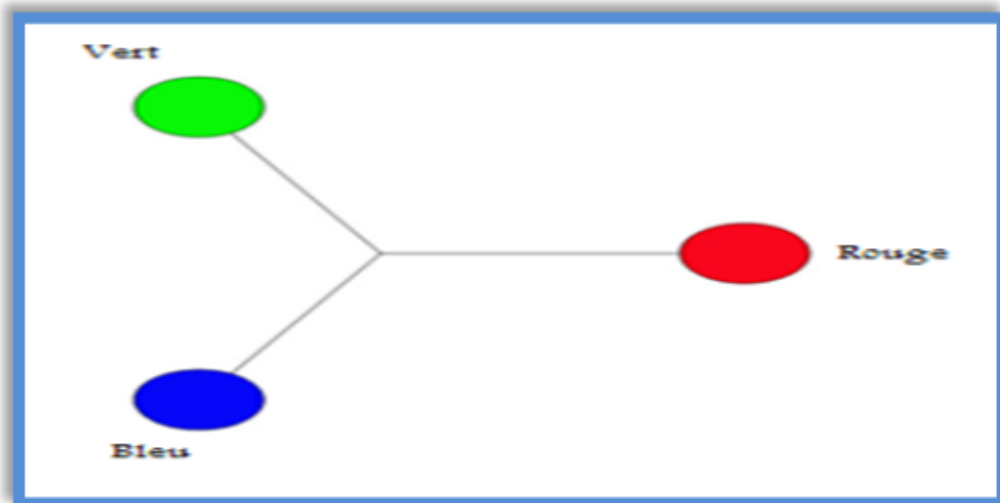


Figure - 4 : Les trois primaires: Rouge, Vert, et Bleu

- **TSL: (Teinte, Saturation, Luminosité):**

Ce modèle peut être représenté sous forme d'un cylindre, mais généralement il est représenté sous forme d'un cône. *La teinte* représente le composant chromatique. *La saturation* se rapporte à la dominance d'une teinte particulière dans une couleur. *La luminosité* d'une couleur se rapporte à l'intensité. [Ben Cheikh.N et al, 11]

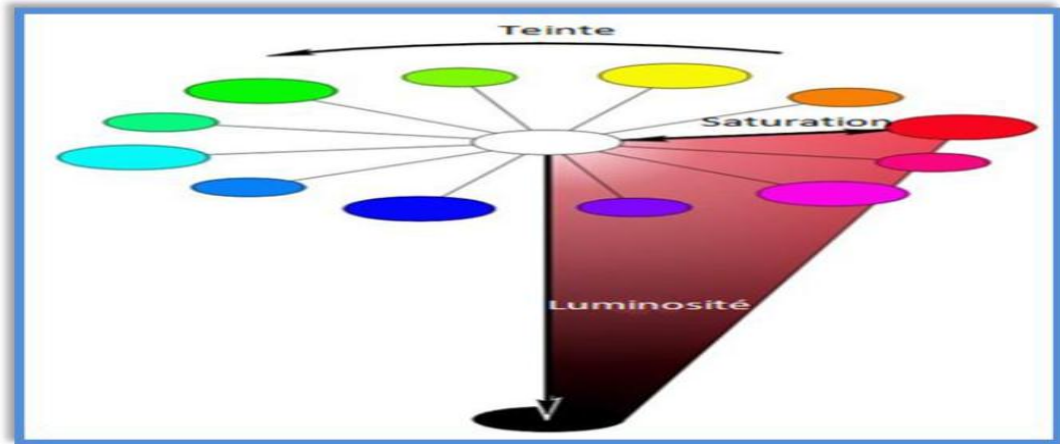


Figure - 5 : TSL (teinte, saturation, luminosité) sous forme cône.

### - L'histogramme

Un histogramme est un graphique statistique permettant de représenter la distribution des intensités des pixels d'une image, c'est-à-dire le nombre de pixels pour chaque intensité lumineuse comme l'exemple suivant:

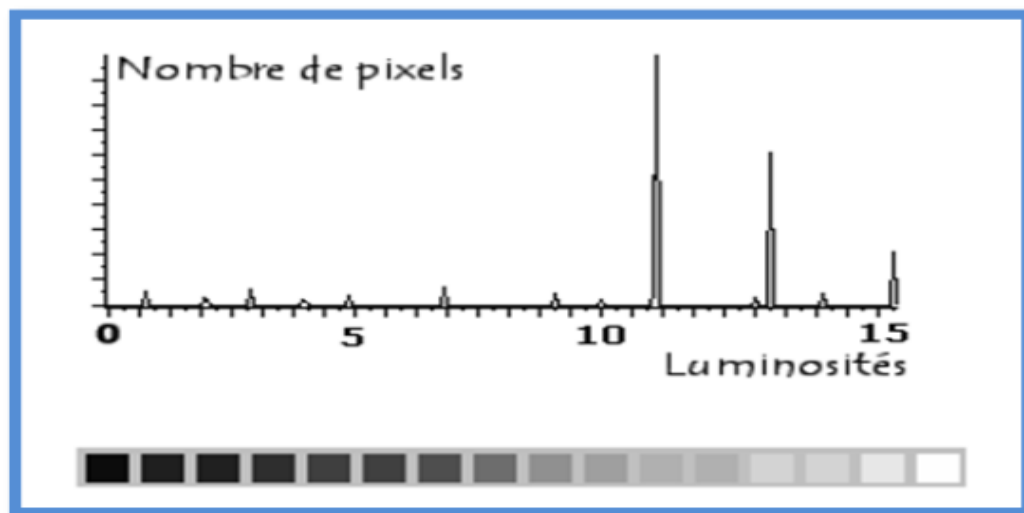


Figure - 6 : Exemple d'un histogramme.

Pour les images en couleur plusieurs histogrammes sont nécessaires. Par exemple pour une image codée en RGB il faut un histogramme représentant la distribution de la luminance et trois d'autres histogrammes représentant respectivement la distribution des valeurs respectives des composantes rouges, bleues et vertes. [Ben Cheikh.N et al, 11]

### 2.2.6. La texture

#### 2.2.6.1. Définition

Bien qu'il n'existe pas de définition théorique de la texture, elle est cependant généralement modélisée comme : "une structure spatiale constituée de l'organisation de primitives ayant chacune un aspect aléatoire.". [Roux. B, 04]

On peut aussi définir une texture comme : "la structure détaillée d'une image, trop fine pour être analysée entièrement, cependant assez grossière pour induire une fluctuation notable des niveaux de gris de pixels voisins ". [Roux. B, 04]

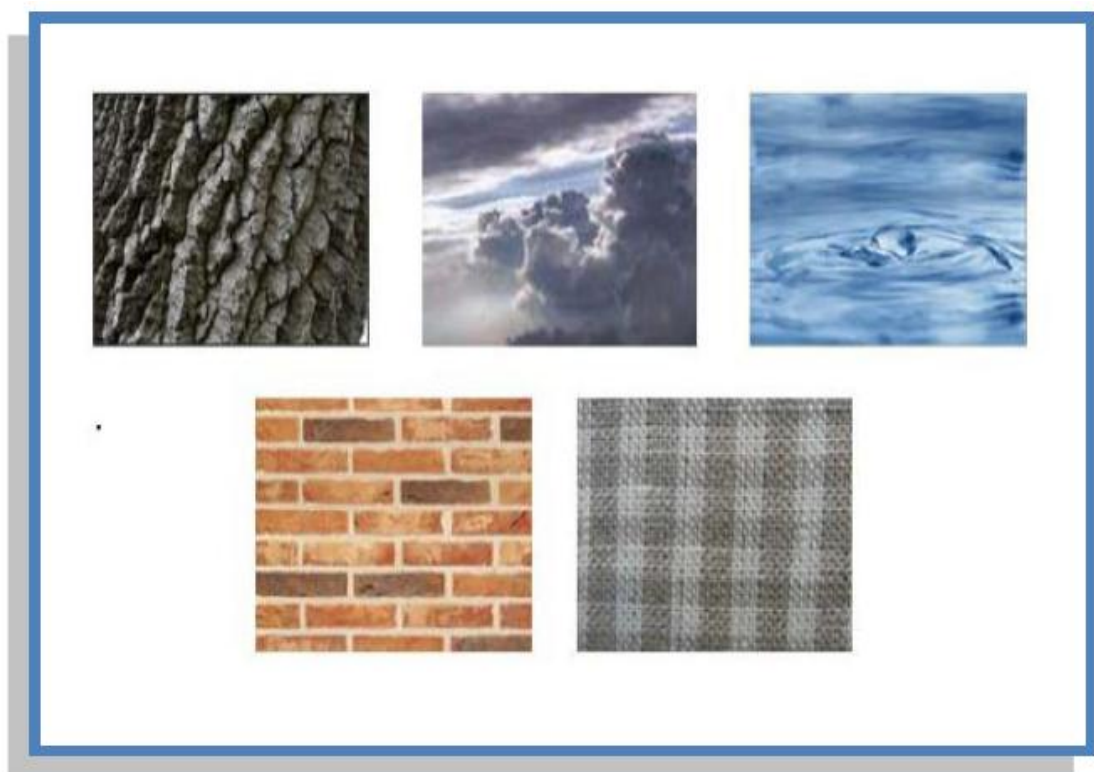


Figure - 7 : Des exemples de textures

#### 2.2.6.2. La matrice de co-occurrence

La méthode de matrice de co-occurrence présente une grande simplicité de mise en œuvre et donne de bons résultats sur la plupart des types d'images. Une matrice de co-occurrence est une matrice de taille  $N \times N$ , où  $N$  est le nombre de niveaux de gris d'une image. Pour un déplacement  $\mathbf{D}$  (translation), un élément  $(i, j)$  de la matrice est défini par le nombre de pixels de l'image de niveau de gris  $j$  situés à  $\mathbf{D}$  d'un pixel de niveau de gris  $i$ . [Iggane. M, 09]

## CHAPITRE II : La segmentation d'images

---

Tous ces indices sont définis pour une valeur de déplacement  $D$ , le choix de ce paramètre est donc très important quant à l'obtention d'un résultat de qualité. [Iggane. M, 09]

La matrice de cooccurrence contient une masse très importante d'informations et est donc difficilement manipulable. De ce fait, quatorze indices qui correspondent à des caractères descriptifs des textures peuvent être calculés à partir de cette matrice. Nous ne présentons ici que neuf de ces indices: [Iggane. M, 09]

$$Mean = \sum_i \sum_j i.C(i, j)$$

$$Variance = \sum_i \sum_j (i - \mu)^2.C(i, j)$$

$$Energy = \sum_i \sum_j C(i, j)^2$$

$$Correlation = \sum_i \sum_j (i - \mu)(j - \mu).C(i, j)$$

$$Entropy = \sum_i \sum_j -C(i, j).log(C(i, j))$$

$$Contrast = \sum_i \sum_j (i - j)^2.C(i, j)$$

$$Homogeneity = \sum_i \sum_j \frac{1}{1 + (i - j)^2}.C(i, j)$$

$$Cluster Shade = \sum_i \sum_j (i + j - 2\mu)^3.C(i, j)$$

$$Cluster Prominence = \sum_i \sum_j (i + j - 2\mu)^4.C(i, j)$$

### 2.2.7. La forme

Au même titre que les caractéristiques de texture, les attributs de forme sont souvent Complémentaires à la description de couleur. Les attributs de forme sont utilisés pour caractériser les objets dans les images. On distingue deux catégories de descripteurs de

formes: les descripteurs basés régions et les descripteurs basés frontières. Les premiers sont utilisés pour caractériser l'intégralité de la forme d'une région. Ces attributs sont robustes aux transformations géométriques comme la translation, la rotation et le changement d'échelle. Les seconds portent sur la caractérisation des contours de la forme. [Roux. B, 04]

### 3. La segmentation

#### 3.1. Définition de la segmentation d'images

Segmenter une image veut dire la subdiviser en parties pour trouver ses contours et ses régions qui sont visuellement distinctes et uniformes par rapport à des propriétés homogènes. Ces régions et contours sont supposés être pertinents, c'est-à-dire que les régions doivent correspondre aux parties significatives des objets du monde réel, et les contours à leurs contours apparents.

La segmentation est un traitement de bas niveau qui consiste à créer une partition de l'image  $A$  en sous-ensembles  $R_i$ , appelés régions tels qu'aucune région ne soit vide, l'intersection entre deux région soit vide et l'ensemble des régions recouvre toute l'image. [Do.M.C, 07]

La segmentation est une étape primordiale dans les systèmes de traitement d'images, elle peut être précéder par les étapes d'acquisition et de prétraitement dans la chaîne du traitement de l'image.

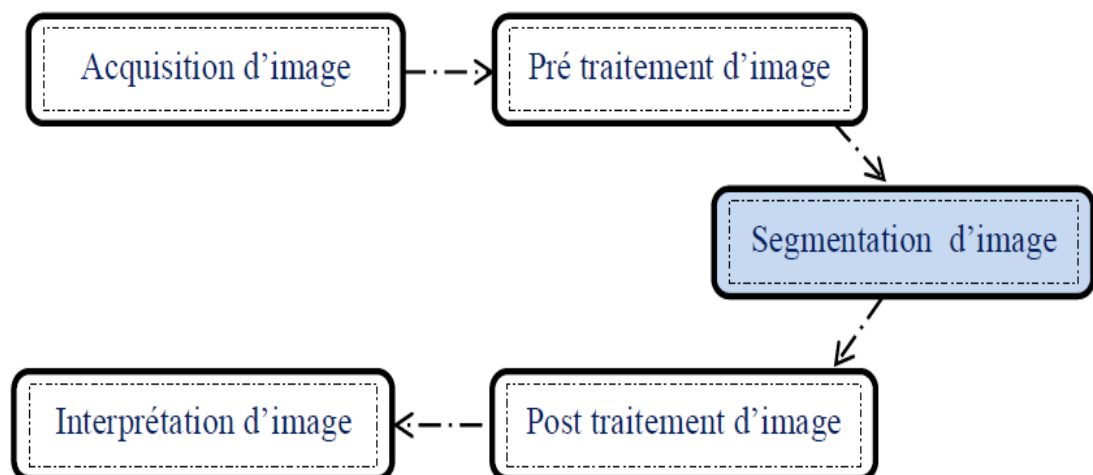


Figure - 8 : Les étapes de traitement d'images.

La phase d'interprétation consiste alors à exploiter le résultat de segmentation afin d'obtenir des informations de haut niveau sur l'image. La qualité des résultats de segmentation influe directement sur les résultats d'interprétation.



### 3.2. Les objectifs de la segmentation

La segmentation d'image n'est pas considérée comme un but en soi, mais dépend du type de traitement fixé a priori. Parmi les objectifs, on cite quelques-unes:

- Extraire (séparer) les entités d'une image pour y appliquer un traitement spécifique
- Segmenter une image pour interpréter son contenu.
- Segmenter une image pour détecter et localiser des objets. [Guediri. K, 12]

### 3.3. Les vues de la segmentation

On peut regrouper les techniques de la segmentation d'images selon les connaissances a priori (segmentation supervisée ou non supervisée), ou selon l'espace d'image (en région ou en contour). [Guediri. K, 12]

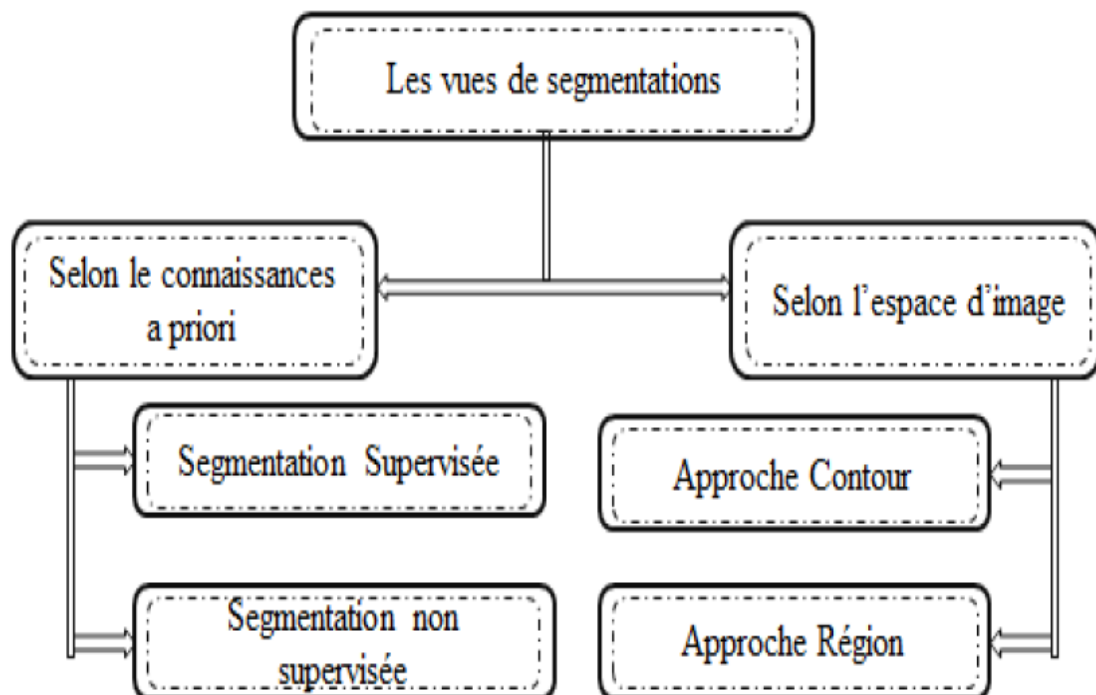


Figure - 9 : les vues de segmentation d'images.

#### 3.3.1. Segmentation selon les connaissances a priori

Il existe deux types :

##### 3.3.1.1 Segmentation supervisée

Ce sont des méthodes dans lesquelles les classes sont connues a priori avant d'effectuer l'opération d'identification des éléments de l'image. Ou on dispose une information pré requise sur l'image à segmenter.

### 3.3.1.2 Segmentation non supervisée

L'intérêt des méthodes non supervisées est qu'elles ne nécessitent aucune base d'apprentissage et par là même aucune tâche préalable d'étiquetage manuel n'est requise. [Guediri. K, 12]

### 3.3.2. Segmentation selon l'espace d'image

Malgré la diversité des techniques de segmentation d'images on peut les classer en deux grandes catégories de segmentation : Les approches de détection de similarité (région), et les approches de détection de discontinuité (contour). [Guediri. K, 12]

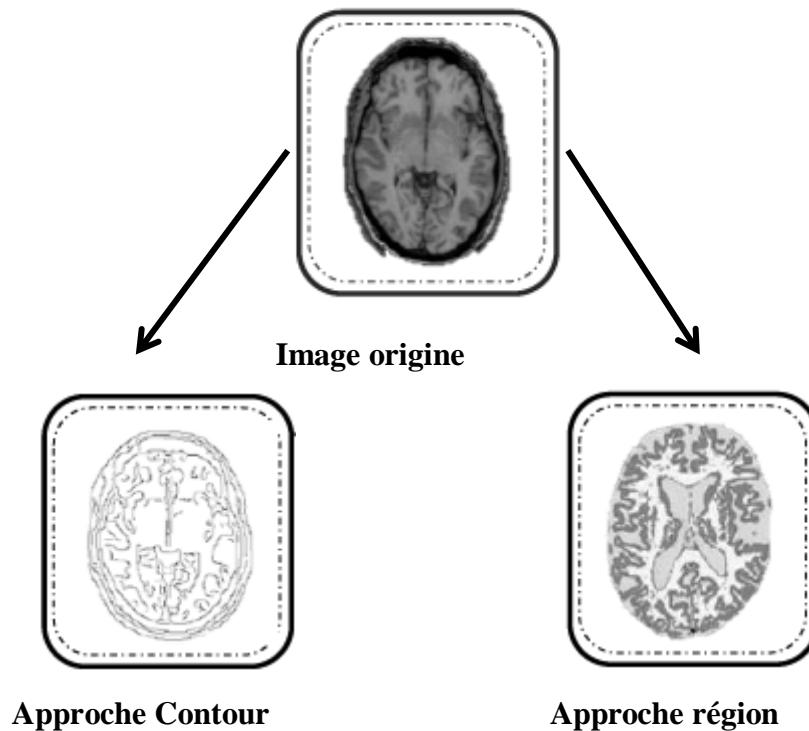


Figure - 10 : Les approches de segmentation selon l'espace d'image.

#### 3.3.2.1. Segmentation par approche contour

La segmentation d'images par l'approche contour (ou frontière) est une méthode permettant de localiser les frontières d'un objet, en détectant les zones de fortes discontinuités de l'image. On considère que les primitives à extraire sont les lignes de contrastes séparant des régions, ou bien des régions de texture différentes.

Dans cette approche, il existe plusieurs méthodes, telles que la méthode de contour actif "snakes" (Serpent) et la méthode de ligne de partage d'eaux "LPE" (watershed).

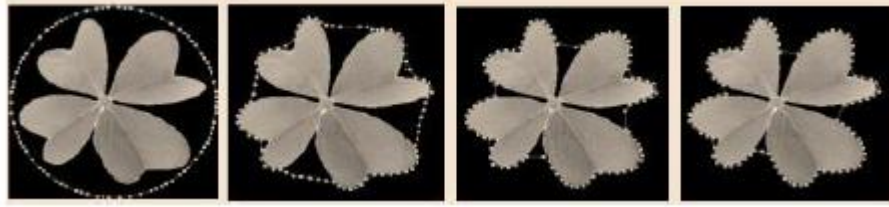


Figure - 11 : Segmentation par la méthode de contour actif

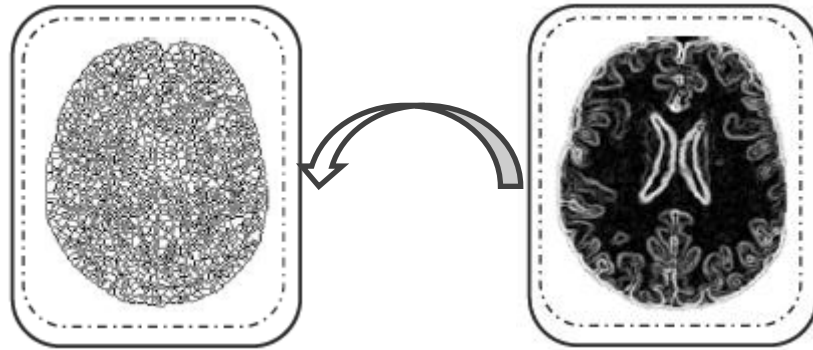


Figure - 12 : Une segmentation obtenue grâce à LPE

### 3.3.2.2. Segmentation par approche région

Les approches régions ont pour but de mettre en évidence les régions homogènes de l'image. Il s'agit de rechercher des ensembles de pixels partageant des propriétés communes. Il existe plusieurs méthodes pour la segmentation dans cette approche, les plus utilisées sont [Guediri. K, 12] : la segmentation par Machines à Support de Vecteurs SVM (Support Vector Machine), la segmentation par Croissance de région (growing region), Segmentations par Division / fusion (Split /Merge), Segmentation par Champ de Markov, Segmentation par classification (K means, Fuzzy-c-means FCM).

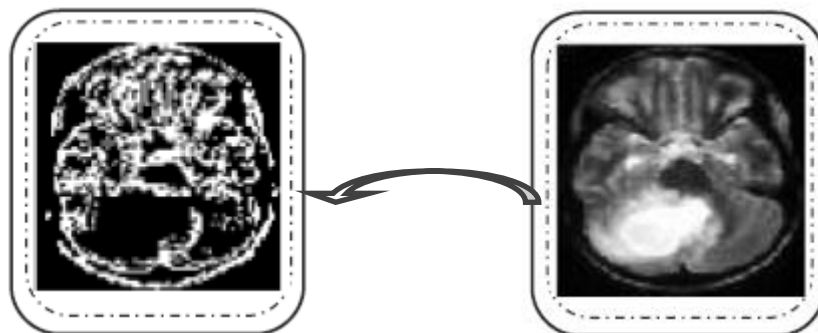


Figure - 13 : La segmentation de type split & merge

Une description détaillée de chacune de ces méthodes est trouvée dans [Jérémy.L et al, 08], [Baillie. J-C, 03],

### **4. Conclusion**

La segmentation d'images est une étape fondamentale dans beaucoup d'applications telle que l'interprétation d'images. Cependant, le choix d'une bonne méthode de segmentation qui possède une haute qualité de performance, parmi un nombre important de méthodes développées dans la littérature, est une tâche difficile.

Dans ce chapitre nous avons donné l'essentiel sur l'image numérique et ses caractéristiques visuelles. Ainsi, le nécessaire pour comprendre le sujet de segmentation d'images. Le but visé était de faciliter au lecteur de comprendre par la suite notre travail qui part d'une segmentation des images numériques et un calcul des caractéristiques visuels.

Dans le prochain chapitre nous allons présenter des définitions, des notions indispensables, des méthodes pour l'interprétation des images et les ontologies afin de comprendre par la suite comment exploiter la segmentation dans ce contexte.

## **CHAPITRE III : Annotation et Ontologie**

- **Introduction**
- **Annotations et Métadonnées**
- **Les ontologies**
- **Conclusion**

### 1. Introduction

Une source d'information, que ce soit sa nature, doit être bien décrite, sinon elle peut demeurer pratiquement inexploitable et impossible de la retrouver. L'annotation est une solution à ce problème. Elle peut être en langage libre ou formel comme les ontologies.

L'utilisation des ontologies pour l'annotation et l'interprétation offre une richesse sémantique et des avantages majeurs.

Le but de ce chapitre est de permettre au lecteur de comprendre le sujet d'annotation et l'utilisation des ontologies dans ce sens. Pour cela, nous consacrons une première section afin de présenter des définitions indispensables pour distinguer entre les termes métadonnée, annotation et annotation sémantique. Ensuite, nous allons montrer les différents types et les différentes méthodes d'annotations.

Dans la deuxième section, nous mettons l'accent sur les ontologies comme modèle de représentation efficace des connaissances d'un domaine. Nous commençons par une définition et des motivations du développement d'une ontologie, ses composantes et son cycle de vie. Ensuite l'ingénierie ontologique et enfin, les outils et les langages qui les supportent.

### 2. Annotations et Métadonnées

#### 2.1. Définitions

##### 2.1.1. Métadonnée

Le terme "méta" vient du grec et qui signifie dénoter quelque chose de nature plus élevée ou plus fondamentale. [Boucetta. Z, 08]

On utilise généralement les méta-données pour parler d'informations descriptives à propos de ressources du Web. Plusieurs définitions sont présentées dans la littérature pour les méta-données. On va exposer quelques-unes:

- La définition des métadonnées la plus simple et consensuelle est la suivante : les métadonnées sont de l'information sur l'information électronique. Toutefois, l'importance prise par les métadonnées aujourd'hui mérite quelques précisions quant à leur définition. Pour **Vellucci** en 1998, les métadonnées sont les informations utilisées pour la description et la

gestion des ressources. Pour **Ercegovac** en 1999, les métadonnées sont des informations pour décrire, identifier et définir une ressource. [**Anis.j, 05**]

- Une description normalisée attachée à une ressource identifiée (sur le Web notamment) [**Boucetta. Z, 08**]

- « Une métadonnée est une donnée sur une donnée » [**Boucetta. Z, 08**].

Une métadonnée est attachée à une ressource identifiée en tant que telle sur le Web et aura plutôt une pertinence à priori et sera plutôt saisie suivant un schéma. Par exemple, le titre et l'auteur d'un document, qui permettront de mettre en place des inférences. [**Boucetta. Z, 08**]

### 2.1.2. Annotation

Plusieurs communautés de chercheurs (informatique, linguistique, rhétorique, etc.) ont étudié l'annotation, et en ont proposé plusieurs définitions.

Selon la communauté des informaticiens plusieurs définitions sont accordées à l'annotation. Nous présentons ci-après les plus célèbres.

- « Une annotation est une information graphique ou textuelle attachée à un document et le plus souvent placée dans ce document ». [**Amourache. F, 08**]

- « un commentaire sur un objet tel que le commentateur veut qu'il soit perceptiblement distinguable de l'objet lui-même et le lecteur l'interprète comme perceptiblement distinguable de l'objet lui-même ». [**Amourache. F, 08**]

- « un commentaire libre situé à l'intérieur de la ressource documentaire ». [**Boucetta. Z, 08**]

### 2.1.3. Annotation sémantique

Dans le cadre du Web, il existe un autre type particulier d'annotation appelée annotation sémantique, c'est à dire interpréter une ressource ou donner sa signification par une description formelle.

Dans la littérature, on trouve plusieurs définitions d'une annotation sémantique, les plus complètes sont :

- « L'annotation sémantique est un processus d'instanciation des données à partir d'une ontologie du domaine ». [**Boucetta. Z, 08**]

- « L'annotation sémantique est définie comme une représentation formelle d'un contenu, exprimée à l'aide de concepts, relations et instances, décrits dans une ontologie, et reliées à la ressource documentaire source ». [Boucetta. Z, 08]

Les annotations sémantiques décrivent le lien entre les entités qui se trouvent dans la ressource d'information (document, image vidéo,...etc) et leurs descriptions sémantiques représentées dans l'ontologie.

Elles permettent ainsi de désambiguïser et d'interpréter les ressources, en faisant référence à une connaissance (habituellement une ontologie) séparée, avec une richesse sémantique considérable. Comme elles servent dans beaucoup de traitements automatiques tels que la classification et la recherche d'informations.

L'annotation sémantique à partir d'ontologie est devenue actuellement l'approche la plus prometteuse pour partager et exploiter l'information sur le Web. [Boucetta. Z, 08]

### 2.2. Méthodes d'annotation

Il existe trois types d'annotation d'images : manuelle, semi-automatique et automatique. La première, comme sa dénomination l'indique, est effectuée manuellement par un humain chargé d'attribuer à chaque image un ensemble de mots clé. L'annotation automatique, quant à elle, est une tâche effectuée par une machine et vise à alléger la charge de l'utilisateur. Le premier type d'annotation augmente la précision et baisse la productivité tandis que le second type baisse la précision de la tâche et augmente sa productivité. Afin de faire un compromis entre ces deux tâches, leur combinaison est devenue nécessaire. C'est ce qui est connu sous le nom "l'annotation semi-automatique". [Hacid. H, 06]

#### 2.2.1. L'annotation manuelle

L'annotation manuelle est entièrement effectuée par un annotateur humain qui place des annotations selon son point de vue.

Alors l'annotation manuelle consiste simplement à mettre en place une interface utilisateur dans laquelle l'utilisateur humain peut sélectionner la ressource à annoter, choisir le modèle formel servant à la création des annotations sémantiques, tout en respectant les contraintes imposées par le modèle formel, créer les annotations voulues sur la ressource sélectionnée. [Amourache. F, 2008]



### 2.2.2. L'annotation semi-automatique

L'annotation semi-automatique, comme son nom l'indique, se devise en une première phase qui s'effectue manuellement et une deuxième qui s'effectue automatiquement.

La phase manuelle, elle consiste à faire une annotation manuelle d'un échantillon d'informations seulement. Puis on procède dans la deuxième phase à une propagation de l'annotation.

L'annotation semi-automatique s'appuie généralement sur un moteur d'extraction d'informations qui est intégré dans un outil d'annotation. Il suggère des annotations à l'utilisateur qui peut ensuite les valider manuellement. [Amourache. F, 08]

Il existe des méthodes qui appliquent un partitionnement ("clustering") des images et des textes associés aux images permettant ainsi de lier le texte et l'image. Avec cette démarche, il est possible de prédire l'étiquette d'une nouvelle image en calculant des probabilités. Minka et Picard en 1995 ont proposé un système semi-automatique d'annotation d'images dans lequel l'utilisateur choisit les régions à annoter dans une image. Une propagation des annotations est effectuée en considérant les textures. Maron et ses collègues en 1998 ont abordé l'annotation automatique mais en utilisant un seul mot clé à la fois. Mori et ses collègues en 1999 ont proposé un "*modèle de co-occurrences*" entre l'image et les mots clés afin de trouver les mots clé les plus pertinents pour une image. L'inconvénient de ce modèle est qu'il nécessite un grand échantillon d'apprentissage pour qu'il soit efficace. Dyugulu et ses collègues en 2002 ont proposé un modèle, appelé "*modèle de translation*", qui est une amélioration du modèle de co-occurrence proposé par Mori et ses collègues en 1999, en utilisant un algorithme d'apprentissage. [Hacid. H, 06]

### 2.2.3. L'annotation automatique

L'annotation automatisée est entièrement réalisée par un outil d'extraction d'informations qui est intégré dans l'outil d'annotation. Le traitement automatique consiste à **Teni** en 2007 :

1. Identifier automatiquement, dans la ressource à annotée, les éléments qui sont pertinents.
2. Déterminer quels sont les concepts de l'ontologie les plus répondus, pour annoter chacun de ces éléments. [Amourache. F, 08]

Pour le texte, l'automatisation se repose sur un apprentissage à partir d'un corpus constitué d'éléments marqués par un expert. Le marquage associe à chaque concept de l'ontologie des éléments du page en rapport avec ce concept. Des mécanismes de raisonnement exploitant l'ontologie sont utilisés pour déterminer le concept le plus spécifique avec lequel un élément doit être annoté. Cependant pour l'image ou la vidéo, il n'existe pas à l'heure actuelle des méthodes automatiques proprement dites. [Amourache. F, 08]

### 2.3. Natures de métadonnées

Différentes natures d'informations peuvent être associées à des ressources:

#### 2.3.1. Métadonnée indépendante du contenu

Une métadonnée liée à la ressource (telle que une image ou une vidéo), mais ne se décrit pas le contenu. Citons à titre d'exemple le nom de l'auteur, date, lieu,...etc

Les métadonnées indépendantes du contenu ne peuvent être extraites de la ressource. [Ben Cheikh. N et al. 11]

#### 2.3.2. Métadonnées dépendantes de contenu

Les données qui se réfèrent directement au contenu peuvent être divisées en deux types :

##### 2.3.2.1. Métadonnées dépendantes de contenu symbolique

C'est métadonnées s'appliquent sur les images et la vidéo. Elles sont liée aux caractéristiques de bas niveau (couleur, texture, forme, .. etc.), et par conséquent appelées des caractéristiques ou des descripteurs visuels. Elles sont faciles à extraire avec un temps de calcul important. On peut extraire des vecteurs de caractéristiques énormes contenant caractéristiques de couleur, de texture, de forme...etc

##### 2.3.2.2. Métadonnées descriptives de contenu sémantique

Les métadonnées sémantiques se réfèrent au contenu sémantique, ou à l'interprétation du sens de la ressource et de sens associée à des signes visuels et des scènes.

L'annotation par des métadonnées descriptives de contenu sémantique, autrement dit c'est l'annotation sémantique, est difficile à automatiser et qui nécessite de nombreux essais pour évaluer la performance des algorithmes d'annotation. Cela due à la subjectivité, c'est-à-dire que chacun peut interpréter une ressource selon son point de vision.

## CHAPITRE III : Annotation et Ontologie

---

Les métadonnées descriptives de contenu peuvent être spécifiées à l'aide de l'une des approches suivantes :

### A. Description à base de texte libre

Pour ce type d'annotation il n'y a aucune structure prédéfinie. L'utilisateur peut annoter en utilisant n'importe quelle combinaison de mots ou de phrases. Cela rend plus facile d'annoter, mais plus difficile à exploiter cette annotation plus tard pour d'autres objectifs tel que la recherche. [Ben Cheikh. N et al. 11]

### B. Description à base des mots clé

Chaque ressource est annotée par une liste de mots-clés qui lui sont associés. Il ya deux possibilités pour choisir les mots-clés : soit L'annotateur utilise des mots clés arbitraires comme l'exige, ou bien il est limité par une liste prédéfinie de mots-clés. [Ben Cheikh. N et al. 11]

### C. Description à base de concepts d'une ontologie

C'est semblable à la description à base de mots-clés. Cependant la différence c'est que, d'une part, ce sont des concepts dans une ontologie. Et d'autre part qu'il ya des relations sémantiques entre les concepts, ce qui permet de enrichir les annotations. [Ben Cheikh. N et al. 11]

L'utilisation d'une ontologie pour faire des annotations descriptives de contenu sémantique, est une solution très efficace. Car une ontologie présente une richesse sémantique décrivant un domaine spécifique sous forme des concepts et relations entre eux. Les annotations sémantiques à base d'ontologie sont développées actuellement autour du web sémantique, et semble l'approche la plus prometteuse pour partager, chercher et exploiter efficacement les ressources d'information. [Ben Cheikh. N et al. 11] Les annotations sémantiques sont:

- Persistantes et Implicites : car elles font référence à une connaissance séparée de la ressource (une ontologie).
- Opérationnelles : car elles sont destinées à être traitées par des machines.

## 3. Les ontologies

L'ontologie est une technologie informatique appartenant au champ de l'ingénierie des connaissances, apparue dans les années 80. Cette technologie a suscité l'intérêt de la communauté scientifique. [Khouri. S, 09]

### 3.1. Définition d'une ontologie

Plusieurs définitions d'ontologies ont été proposées. Nous retenons ici la définition suivante : «*C'est une représentation formelle, explicite, référençable et consensuelle de l'ensemble des concepts partagés d'un domaine sous forme de classes, de propriétés et de relations qui les lient* ». Les termes les plus importants dans cette définition sont donc :

- **Formelle** : l'ontologie est définie dans un langage traitable par la machine.
- **Explicite** : l'ensemble des concepts et propriétés d'une ontologie sont spécifiés explicitement.
- **Référençable** : signifie que tout concept de l'ontologie peut être référencé de manière unique afin d'explicitier la sémantique de l'élément référençant.
- **Consensuelle** : signifie que l'ontologie est admise et acceptée par l'ensemble des acteurs d'une communauté. [Khouri. S, 09]

Plusieurs autres définitions du concept ontologie ont été proposées. Ces définitions sont souvent des raffinements de définitions déjà proposées et/ou sont complémentaires avec elles.

### 3.2. Les motivations du développement des ontologies

Le domaine de l'ontologie attire l'attention parce qu'une ontologie fournit :

- **Le partage sémantique** : C'est l'une des premières motivations. La sémantique représente l'interprétation selon un point de vue ou un contexte particulier. Le partage sémantique doit donc être guidé par le contexte et le but à atteindre. Cette compréhension commune du domaine s'effectue au travers d'une ontologie. Ainsi, une ontologie fournit un vocabulaire partagé pour une compréhension commune du domaine traité. Par exemple, en médecine, les termes scientifiques sont désignés par des abrégés. Les mêmes abrégés peuvent se trouver dans plusieurs services mais ayant des sens différents; par conséquent leur utilisation peut prêter à une confusion ; un diagnostic automatique dans ce domaine est impossible sauf que si la compréhension soit commune. Le partage et la communication doivent être présents entre les personnes, entre les personnes et les agents logiciels et enfin entre les agents logiciels eux même. [Ben Cheikh. N et al, 11]

Les ontologies offrent une structuration et une sémantique facilitant considérablement la fourniture d'un service.

- **L'explicitation des spécifications des systèmes:** La plupart des logiciels conventionnels sont construits avec une conceptualisation implicite et que la nouvelle génération des systèmes utilisant les travaux en intelligence artificielle devrait être basée sur une conceptualisation explicitement représentée. En effet, l'ontologie fournit une classification des objets que doit manipuler le système. [Ben Cheikh. N et al, 11]

- **L'analyse du savoir sur un domaine:** la spécification des termes d'un domaine est très importante; c'est ce qui permet d'analyser le savoir sur ce domaine. Ce savoir peut, par la suite, être utilisé dans différentes applications. [Ben Cheikh. N et al, 11]

- **L'indexation et la recherche d'informations :** Dans le Web Sémantique, les ontologies sont utilisées pour déterminer les index conceptuels décrivant les ressources sur le Web. [Ben Cheikh. N et al, 11]

### 3.3. Les composantes des ontologies

Une ontologie se compose des éléments suivants : les **concepts** d'un domaine, modélisés par les **classes** et leurs **attributs**, les **relations** entre ces concepts, et les **axiomes**. [Khouri. S, 09]

#### 3.3.1. Les concepts

Un concept peut se définir comme une entité composée de trois éléments distincts :

- Le **terme** : exprimant le concept en langage naturel.
- **Notion** ou **intension** du concept : la signification du concept.
- Les **objets** dénotés par le concept, appelés également « réalisations » ou « extensions » du concept.

Prenons l'exemple du concept VOITURE. Le terme de ce concept est le nom commun voiture. Sa notion est d'être un véhicule de transport à roues. Son extension est l'ensemble des marques de voitures existantes (Renault, Kia ...). [Khouri. S, 09]

#### 3.3.2. Les classes

Elles représentent le centre d'intérêt de l'ontologie et décrivent les concepts d'un domaine. Une classe peut avoir des sous-classes qui représentent des concepts plus spécifiques que la super classe (ou classe supérieure). [Khouri. S, 09]

Une classe peut avoir des instances. Ces instances sont des entités réelles de cette classe. Elles sont une représentation des extensions du concept. Par exemple : Renault est une instance de la classe Voiture. Il est à noter qu'une ontologie ainsi que l'ensemble des instances de toutes ses classes constituent une base de connaissances. [Khouri. S, 09]

### 3.3.3. Les attributs

Les attributs décrivent les propriétés des classes et des instances de l'ontologie. [Khouri. S, 09]

### 3.3.4. Les relations

Désignent les associations existantes entre les concepts de l'ontologie. La relation principale dans une ontologie est la relation d'hierarchie ou de subsomption « is-a » ou « est-un » (sous classe d'une classe), elle fournit une taxinomie de l'ontologie. Ainsi, il peut exister d'autres types de relations sémantiques entre les concepts.

Il est important de signaler que la richesse sémantique d'une ontologie dépend fortement à la définition des relations sémantiques.

### 3.3.5. Les axiomes

Désignent les assertions acceptées comme vraies dans le domaine étudié. Les axiomes et les règles permettent de vérifier la cohérence d'une ontologie, et aussi d'inférer de nouvelles connaissances. Exemple : « Si deux personnes sont frères, alors il existe quelqu'un qui est la mère de chacun d'eux ». [Khouri. S, 09]

## 3.4. Le cycle de vie des ontologies

Etant donné que les ontologies sont destinées à être utilisées comme des composants logiciels dans des systèmes informatiques répondant à des objectifs opérationnels différents, leur développement doit s'appuyer sur les mêmes principes que ceux appliqués en génie logiciel. En particulier, elles doivent être considérées comme des objets techniques évolutifs et posséder un cycle de vie spécifique. [Ben Cheikh. N et al, 11]

Ce cycle de vie comprend une étape initiale d'évaluation des besoins, une étape de construction, une étape de diffusion, et une étape d'utilisation. Après, chaque utilisation significative, l'ontologie et les besoins sont réévalués et l'ontologie peut être étendue et, si nécessaire, en partie reconstruite. [Chergui. N, 08]

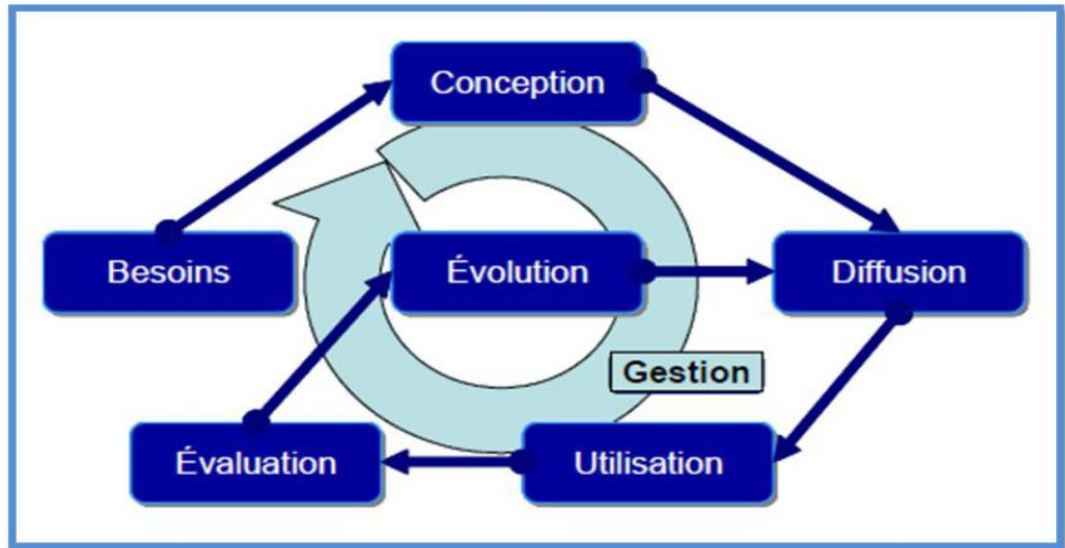


Figure - 14 : Cycle de vie d'ontologie.

Le cycle de vie par évolution de prototypes permet à l'ontologiste de retourner de n'importe quel état à n'importe quel autre si une certaine définition manque ou est erronée. Ainsi, ce cycle de vie permet l'inclusion, le déplacement ou la modification de définitions n'importe quand durant ce cycle. L'acquisition, la documentation et l'évaluation de connaissances sont des activités de support qui sont effectuées pendant la majorité de ces états.

### 3.5. L'ingénierie ontologique

L'ingénierie ontologique peut être définie comme un domaine de recherche visant à proposer des aspects pratiques, essentiellement des méthodes, des outils et des langages dédiés à l'application des résultats de la théorie des ontologies dans la construction d'ontologie. [Mehanna. M et al, 07]

A l'heure actuelle, il n'y a pas encore un consensus sur les meilleures pratiques à suivre lors du développement d'une ontologie. Il existe plus de 33 méthodes d'ingénierie ontologique. Ces méthodologies peuvent porter sur l'ensemble du processus et guider l'ontologiste à toutes les étapes de la construction. [Ben Cheikh. N et al, 11]

Cependant, quel que soit la méthode adoptée, celle-ci doit obligatoirement se référer à des règles fondamentales dans la conception des ontologies qui sont:

- Il n'y a pas qu'une seule façon correcte pour modéliser un domaine, il y a toujours des alternatives viables. La meilleure solution dépend presque toujours de l'application que nous voulons mettre en place et des évolutions que nous anticipons.

- Le développement d'une ontologie est nécessairement un processus itératif. Ce qui consiste à aborder l'ontologie de manière globale et d'y revenir la raffiner et la compléter. Le processus itératif se poursuivra tout le long du cycle de vie de l'ontologie.

- Les concepts dans une ontologie doivent être très proches des objets (physiques ou logiques) et des relations dans le domaine d'intérêt. Ces concepts doivent refléter un modèle tel qu'il est dans la réalité.

Généralement, le processus de construction d'une ontologie est une collaboration qui réunit des experts du domaine de connaissance, des ingénieurs de la connaissance, voire les futurs utilisateurs de l'ontologie. Cette collaboration ne peut être fructueuse que si les objectifs du processus ont été clairement définis, ainsi que les besoins qui en découlent. [Frédéric. F, 02]. Ce processus doit comporter les étapes suivantes :

### 3.5.1. Spécification des besoins

La construction d'une ontologie commence par la définition d'un domaine et de sa portée. C'est à dire, il faut trouver des réponses à des questions comme [Pham, S. 04] :

- Quel est le domaine que l'ontologie couvrira?
- Quels sont les buts de l'utilisation de l'ontologie?
- A quels types de questions l'information contenue dans l'ontologie devra-t-elle fournir des réponses?
- Qui va utiliser et maintenir l'ontologie?

Une fois le but défini, le processus de construction de l'ontologie peut démarrer, en commençant par la phase de conceptualisation.

### 3.5.2. Conceptualisation

L'objectif est d'organiser et de structurer la connaissance acquise durant l'étape de spécification en utilisant des représentations externes qui sont indépendantes des paradigmes de représentation de connaissances et des langages d'implémentation dans lesquels l'ontologie va être formalisée et implémentée. L'idée est de combler graduellement le canal entre les moyens d'expressions des intéressés et les langages d'implantation des ontologies. Les représentations intermédiaires utilisées sont : les taxonomies de concepts, les diagrammes des relations binaires, le glossaire des termes, le dictionnaire des concepts, le tableau des relations binaires, spécifier des contraintes sur les attributs dans une table d'attributs, spécifier des axiomes sur les concepts dans une table d'axiomes logiques, décrire les instances des concepts dans une table d'instances.



### 3.5.3. Formalisation

Une fois le modèle conceptuel structuré, il faut le traduire dans un formalisme formel. Grâce à la formalisation, les définitions des concepts sont plus explicites et précises. L'objectif est de faciliter l'interprétation de l'ontologie.

Parmi les nombreux langages de formalisation des ontologies, il y a trois grandes familles: les langages à base de frames, les modèles des graphes contextuels et les logiques de description.

### 3.5.4. Implémentation

Finalement, on doit implémenter l'ontologie dans un langage. Le langage choisit doit correspondre avec le modèle de formalisation. Une introduction brève des langages d'ontologie est présentée dans la session suivante. [Pham. S, 04]

## 3.6. Les langages des ontologies

Différents langages de spécification d'ontologies issus des formalismes précédemment présentés sont apparus à partir des années 1990, tels que CycL et KIF, LOOM, F-Logie et OCML. [Chergui. N, 08]

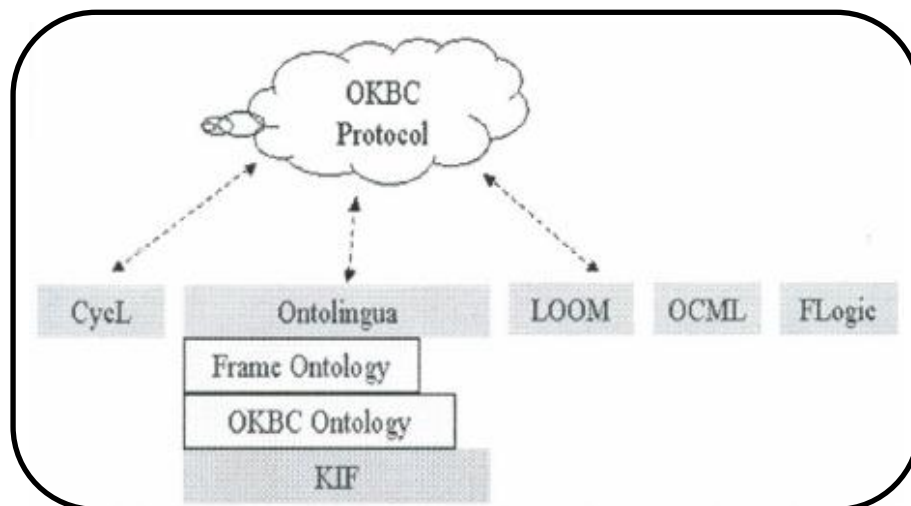


Figure - 15 : Langage traditionnel d'ontologies.

L'explosion des technologies d'Internet a mené à la création des langages pour l'exploitation des caractéristiques du Web, ces langages sont appelés généralement les langages basés Web ou les langages d'annotation d'ontologies, leurs syntaxe est basée sur l'existence d'annotation comme HTML et XML. [Chergui. N, 08]

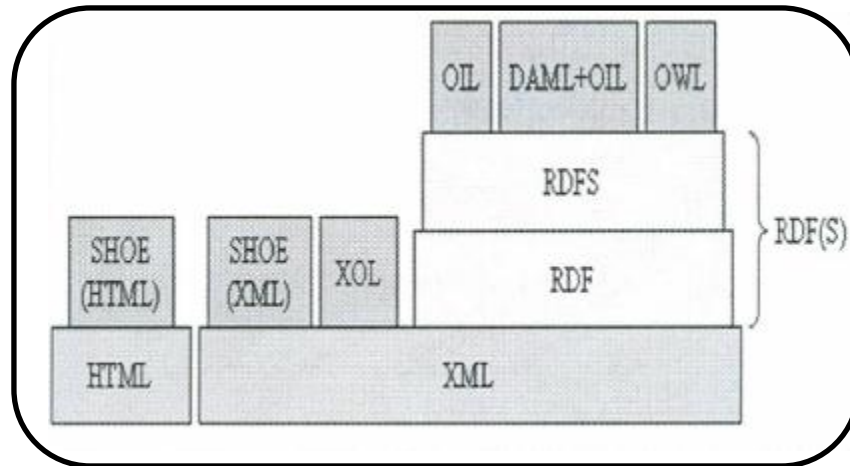


Figure - 16 : Langages d'ontologies

- **SHOE** (Simple HTML Ontology Extensions) : en 1996 est le premier langage d'annotation d'ontologies. Ce langage combine les Frames et les règles de production ce qui lui permet de représenter des concepts, des taxonomies, des relations, et aussi des règles. Ces dernières permettent d'inférer de nouvelles connaissances. [Chergui. N, 08]

- **XML** en 1998, qui a été très vite adopté comme un standard pour les échanges d'informations sur le Web par le W3C (World Wide Web Consortium), **SHOE** a été modifié de telle sorte qu'il puisse supporter des documents structurés décrits en XML.

D'autres langages ont par la suite été créés sur la base de la syntaxe de XML.[Chergui. N, 08]

- **RDF** (Ressource Description Framework) : il a été développé par W3C comme un langage basé sur les réseaux sémantiques pour décrire les ressources du Web

[Chergui. N, 08]. RDF est un langage d'encodage de la connaissance sur les pages

Web. RDF permet de décrire des choses simplement et sans ambiguïté. Toute chose (ressource) est décrite par des phrases minimales, composées d'un sujet, d'un verbe et d'un complément, on parle alors de déclaration *RDF*. [Ben Cheikh. N et al, 11]

- **OWL** est un langage de description d'ontologie conçu pour la publication et le partage des ontologies sur le web sémantique. Il définit un vocabulaire riche pour la description des ontologies complexes. OWL est basé sur une sémantique formelle définie par une syntaxe rigoureuse. OWL peut être utilisé pour représenter explicitement les sens des termes de vocabulaire et les relations entre ces termes.

Le langage OWL se compose de trois sous langages qui proposent une expressivité croissante, chacun conçu pour des communautés de développeurs et des utilisateurs spécifiques: OWLLite, OWLDLet OWLFull. [Ben Cheikh. N et al, 11]

### 3.7. Les environnements et des outils de développement d'ontologies

Les éditeurs d'ontologie constituent des outils facilitant à la construction d'ontologies.

Il existe différents éditeurs d'ontologie; les plus connus sont:

#### 3.7.1. Ontolingua

ONTOLINGUA développé à l'Université de *Stanforâ*, le serveur *Ontolingua* est le plus connu des environnements de construction d'ontologies en langage *Ontolingua*. Il consiste en un ensemble d'outils et de services qui supportent la construction en coopération d'ontologies, entre des groupes séparés géographiquement. [Mehanna. M et al, 07]

Il ya trois différentes possibilités d'intégrer les ontologies Ontolingua

- ✓ Inclusion: Une ontologie inclut et utilise les définitions d'autres ontologies.
- ✓ Restriction: l'ontologie importe les définitions depuis d'autres ontologies et les rend plus spécifiques.
- ✓ Raffinement polymorphe on redéfinit une définition importée depuis n'importe quelle ontologie.

#### 3.7.2. Ontosaurus

Logiciel développé à l'Institut des Sciences de l'Information à l'Université de Southern California. **Ontosaurus** consiste en un serveur utilisant **LOOM** comme langage de représentation des connaissances et en un serveur de navigation réalisant dynamiquement des pages HTML qui affichent la hiérarchie de l'ontologie; le serveur utilise des formulaires HTML pour permettre à l'utilisateur d'éditer l'ontologie. [Ben Cheikh. N et al, 11]

#### 3.7.3. ODE

ODE (ontology design environment) est développée au laboratoire d'Intelligence Artificielle de l'Université de *Madrid*, permet de construire des ontologies au niveau connaissance. [Chergui. N, 08].

#### 3.7.4. Protégé

PROTEGE-OWL [protégé] est une interface modulaire, développée au stanford médical informatique de l'université de stanford<sup>7</sup>, permettant l'édition, la visualisation, le contrôle (vérification des contraintes) d'ontologie, l'extraction d'ontologie à partir de sources textuelles, et la fusion semi –automatique d'ontologies, Le modèle de connaissance de PROTEGE-OWL est issu du modèle des frames et contient des classes (concepts), des slots (propriétés) et des facettes (valeurs des propriétés et contraintes), ainsi que des instances des classes et des propriétés . PROTEGE-OWL autorise la définition de méta-classes, dont les instances sont des classes ce qui permet de créer son propre modèle de connaissances avant

## CHAPITRE III : Annotation et Ontologie

de bâtir une ontologie. De nombreux plug-ins sont disponibles ou peuvent être ajoutés par l'utilisateur.

L'interface est très bien conçue, et l'architecture logicielle permettant l'insertion de plug-ins pouvant apporter de nouvelles fonctionnalités (par exemple, la possibilité d'importer et d'exporter les ontologies construites dans divers langages opérationnels de représentation tels que OWL ou encore la spécification d'axiomes) ont participé au succès de PROTEGE-OWL qui regroupe une communauté d'utilisateurs très importantes et constitue une référence pour beaucoup d'autres outils. Grâce à toutes ces spécifications que dispose PROTEGE-OWL, notre choix portera sur cet outil pour faciliter l'implémentation de l'ontologie d'application. Il faut mentionner qu'il est gratuit.

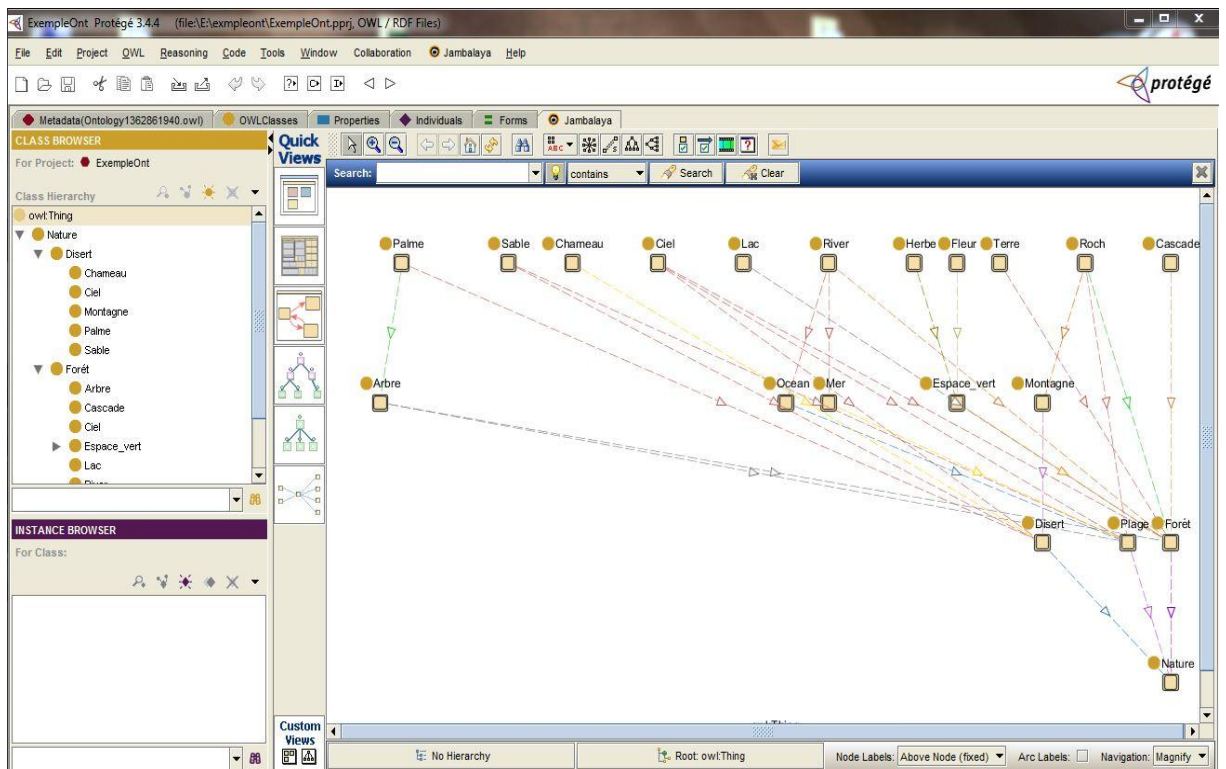


Figure - 17 : l'interface de Protégé

### 4. Conclusion

Les annotations sémantiques à base d'ontologie décrivent les entités qui se trouvent dans les ressources d'informations (document, image vidéo,...etc) et le lien entre ces entités en se basant sur les concepts et les relations dans cette ontologie.

Elles permettent ainsi de désambiguïser et d'interpréter les ressources. Comme elles servent dans beaucoup de traitements automatiques tels que: la création de plusieurs niveaux d'abstractions des ressources, la classification et la recherche d'informations.

L'annotation sémantique à partir d'ontologie est devenue actuellement l'approche la plus prometteuse pour partager et exploiter l'information sur le Web.

Dans le chapitre suivant, nous allons présenter l'approche conceptuelle que nous avons suivie pour atteindre l'objectif de notre travail

## **CHAPITRE IV : Conception**

- **Introduction**
- **Conception de notre ontologie**
- **Conception de notre système d'annotation  
sémantique semi-automatique**
- **Conclusion**

## 1. Introduction

Après avoir exposé les différents axes requis pour la compréhension du contexte de notre travail, dans ce chapitre nous allons présenter toutes les étapes de conception de notre système d'annotation sémantique semi-automatique des images.

Nous allons commencer par la conception de notre ontologie *Onto\_Nature* sur le domaine de la nature.

Les sections qui suivent seront consacrées à la description conceptuelle de notre système d'annotation sémantique semi-automatique des images, qui va exploiter notre ontologie pour interpréter les images.

## 2. Conception de notre ontologie « Onto\_Nature »

Avent de procéder une annotation sémantique il faut avoir une model sémantique du domaine en considération, telle qu'une ontologie. Pour cela, la première étape dans notre travail est la création d'une ontologie.

Il existe une multitude de méthodes d'ingénierie ontologique. Cependant, il n'y pas de consensus sur les principes qui doivent guider la modélisation ontologique. La plupart de ces méthodes visent à identifier les concepts et les relations à partir de documents du domaine, ou de questions posées aux experts.

Comme nous avons présenté dans le troisième chapitre sur les ontologies, le processus de construction d'une ontologie s'articule principalement sur les suivantes:

### Phase 1: Spécification des besoins

La construction d'une ontologie commence par la définition d'un domaine et de sa portée. Pour cela :

- Nous avons choisi le domaine de la nature comme domaine pour notre ontologie.
- Les buts de l'utilisation de l'ontologie sont :
  - Faire l'annotation sémantique d'une base d'image sur ce domaine
  - Enrichir le résultat de l'annotation par d'autres concepts de niveaux d'abstraction plus élevé par l'exploitation des relations sémantiques dans l'ontologie.
  - Servir d'autres travaux tels que la recherche d'images.

## CHAPITRE IV : Conception

---

- L'ontologie sert principalement à fournir des concepts avec lesquels on va faire l'annotation. De plus, ses relations vont servir pour avoir plusieurs niveaux d'abstraction dans les annotations des images.

- Les utilisateurs potentiels de l'ontologie sont : les personnes chargés de processus d'annotation. Ainsi que ceux responsables de développer un mécanisme pour la recherche sémantique des images

### **Phase 2 : Conceptualisation**

Puisque notre ontologie est sur le domaine de la nature, nous avons sélectionné quelques concepts visuels de la nature autour du désert, la plage et le forêt. Puis on a suivi les étapes suivantes :

**Etape 1** : énumérer et identifier les termes important de l'ontologie (dictionnaire de glossaire).

L'étude faite sur le domaine de la nature nous a permis de dégager une liste de termes tels que: forêt, désert, plage, Arbre, cascade... etc

**Etape 2** : définir les relations entre les concepts.

Dans notre ontologie nous sommes intéressés par la hiérarchie entre les classes. Pour cela, nous avons utilisé la relation *partie-de* (*part- of*)

<i>Concept</i>	<i>Concept mère</i>
<b>Nature</b>	/
<b>Plage</b>	<b>Nature</b>
<b>Forêt</b>	<b>Nature</b>
<b>Désert</b>	<b>Nature</b>
<b>Rivers</b>	<b>Forêt</b>
<b>Rocher</b>	<b>Plage, Forêt, Désert</b>
<b>Cascade</b>	<b>Forêt</b>
<b>Sable</b>	<b>Désert</b>



<b>Mer</b>	<b>Plage</b>
<b>Herbe</b>	<b>Espace vert</b>
<b>Chameau</b>	<b>Désert</b>
<b>Fleurs</b>	<b>Espace vert</b>
<b>Espace vert</b>	<b>Forêt</b>
<b>Ciel</b>	<b>Forêt, Plage, Désert</b>
<b>Arber</b>	<b>Forêt, Plage</b>
<b>Montagne</b>	<b>Désert</b>
<b>Palmier</b>	<b>Désert</b>

**Tableau 1 : Classes et hiérarchie des classes de l'ontologie**

Les schémas suivants représentent graphiquement cette taxinomie.

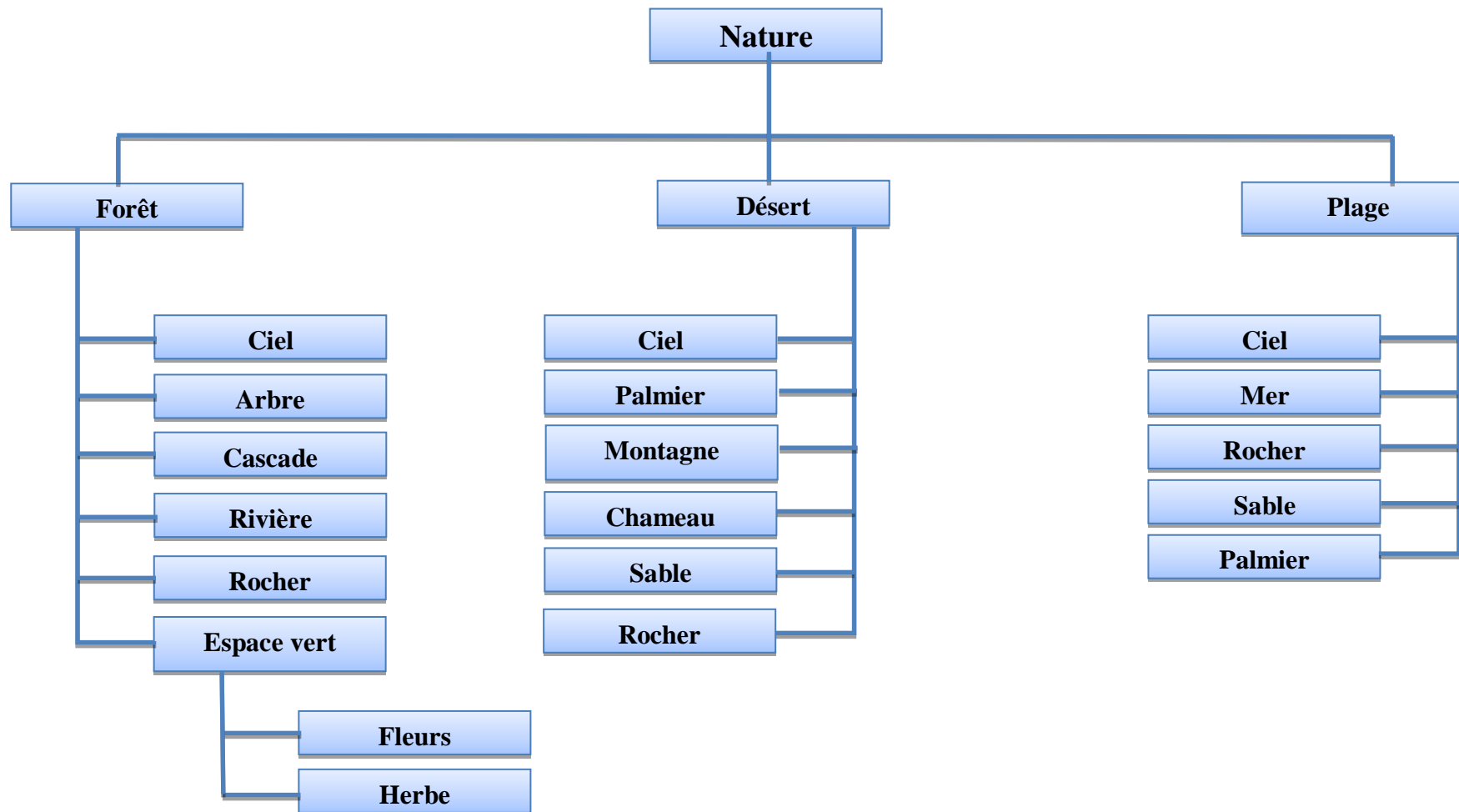


Figure - 18 : Hiérarchie de notre ontologie « Onto\_Nature »

### 3. Conception de notre système d'annotation sémantique semi-automatique

Comme nous avons illustré dans le chapitre précédent, il existe trois approches pour procéder une annotation:

- Approche d'annotation manuelle : consiste à annoter chacune des images de la base manuellement par les concepts les plus apparents dans cette image, et d'associer à chaque concept un poids.
- Approche d'annotation Semi-automatique : consiste à effectuer une partie du travail manuellement puis de continuer d'une façon automatique **par propagation de concepts et/ou par plusieurs niveaux d'abstraction.**
- Approche d'annotation Automatique : aucun travail effectué dans ce sens.

Notre but principal consiste à développer un système d'annotation sémantique semi-automatique, par l'exploitation de notre ontologie *Onto\_Nature*. Il se compose des modules suivants:

- Un module de filtrage des images.
- Un module de segmentation et extraction des caractéristiques des images.
- Un module d'annotation sémantique à base d'ontologie.

La figure ci- dessous illustre l'architecture générale de notre système :

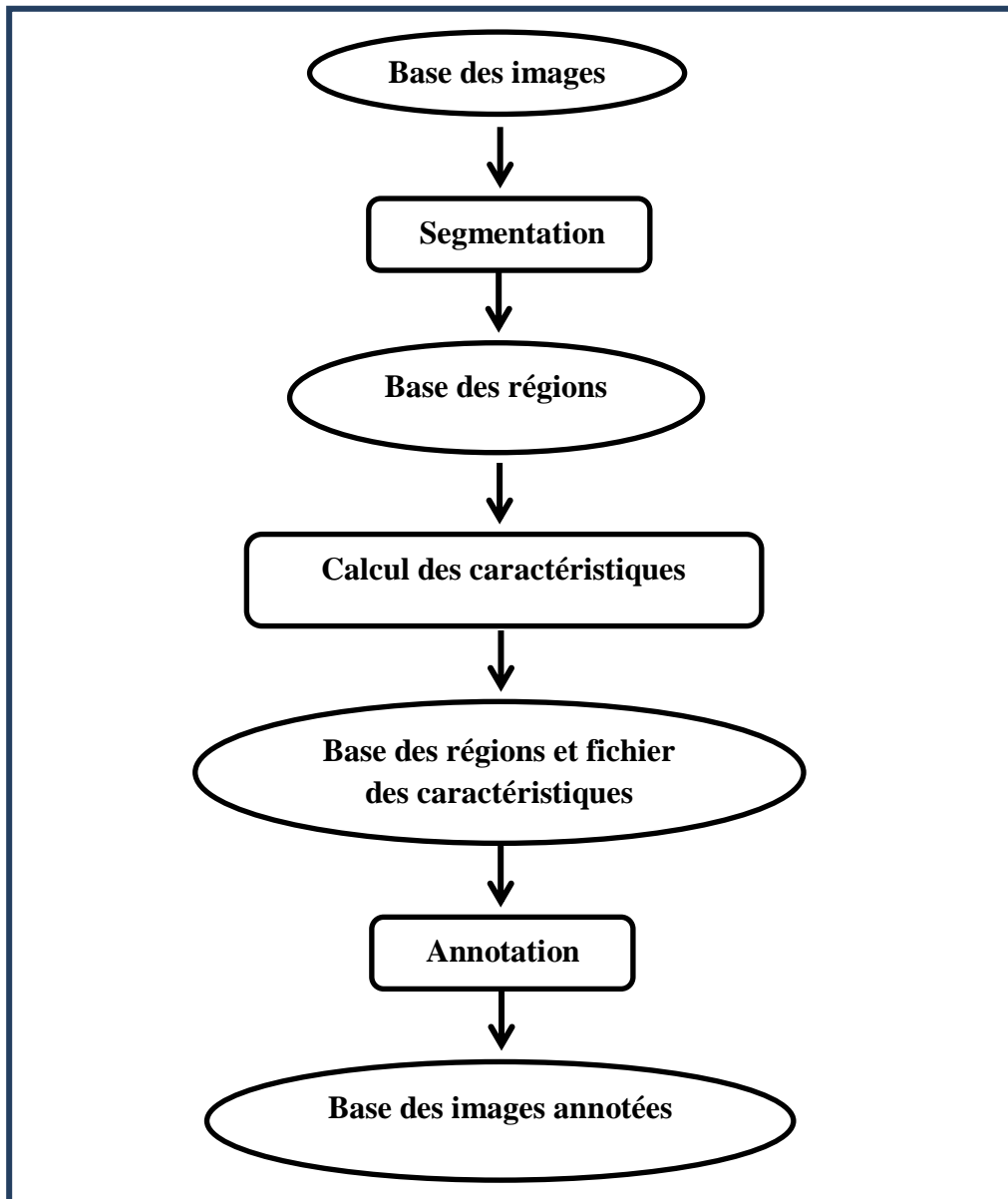


Figure - 19 : Architecture générale de notre système.

### 3.1. Le choix d'une base d'images

Comme nous avons indiqué précédemment, nous avons choisi quelques concepts de la nature pour notre ontologie. Par conséquent, nous allons construire une base d'images reflétant ces concepts.

### 3.2. La segmentation de la base d'images

La segmentation s'applique pour chaque image de la base de données. Elle consiste à extraire les régions de l'image. On obtient comme résultat une base de toutes les régions des images.

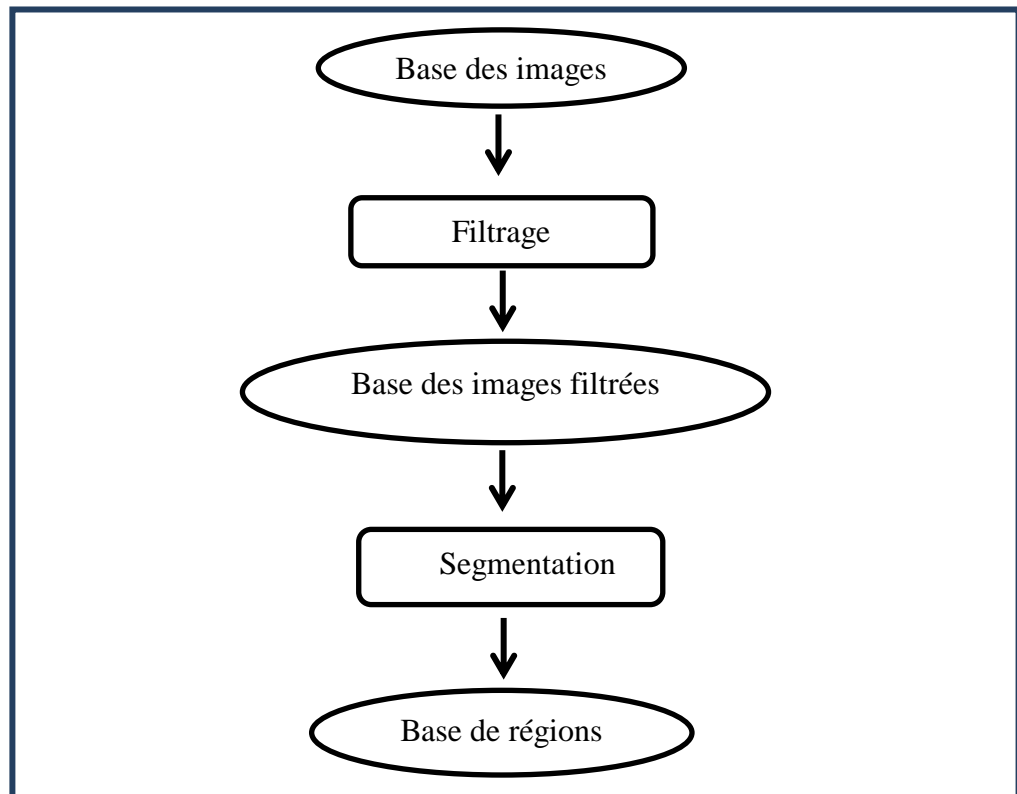
## CHAPITRE IV : Conception

---

Chaque région est enregistrée sous forme d'une image avec un fond blanc.

A cet effet, et pour calculer par la suite les caractéristiques visuels, on parcourt toute la matrice de pixels. Et donc, les pixels d'arrière-plan sont pris en considération. Ce qui va erroné les résultats de calcul.

Pour éviter ce genre de problème, et ne pas prendre en considération le fond d'une région, nous allons effectuer un filtrage.



**Figure - 20 : Architecture de module de segmentation**

### 3.2.1. Le filtrage des images

Le but de filtrage est de distinguer entre les pixels blancs dans une région de celles de l'arrière-plan. Pour cela, nous allons diminuer la couleur d'un pixel dans l'image par un degré, c'est-à-dire sa valeur RGB (255, 255, 255) devienne RGB (254, 254, 254). Par conséquent, lors de parcourt de la matrice, tout pixel dont la valeur est (255, 255, 255) sera ignoré.

### 3.2.2. La segmentation des images filtrées

Comme nous avons présenté dans le premier chapitre sur la segmentation d'images, il existe une panoplie de méthodes de segmentation. Cependant, notre objectif principal n'était pas de choisir un Algorithme de segmentation, mais de travailler sur des régions pour faire les annotations. Pour cela, et pour ne pas tarder dans cette étape, nous allons choisir de la procéder d'une façon manuelle.

Cette phase consiste à segmenter chacune des images de la BD en des régions.

### 3.3. L'extraction des caractéristiques des régions

Ce module consiste à calculer pour chaque région ses caractéristiques visuelles: couleur et texture. On obtient à la fin un fichier contient les régions avec les caractéristiques.

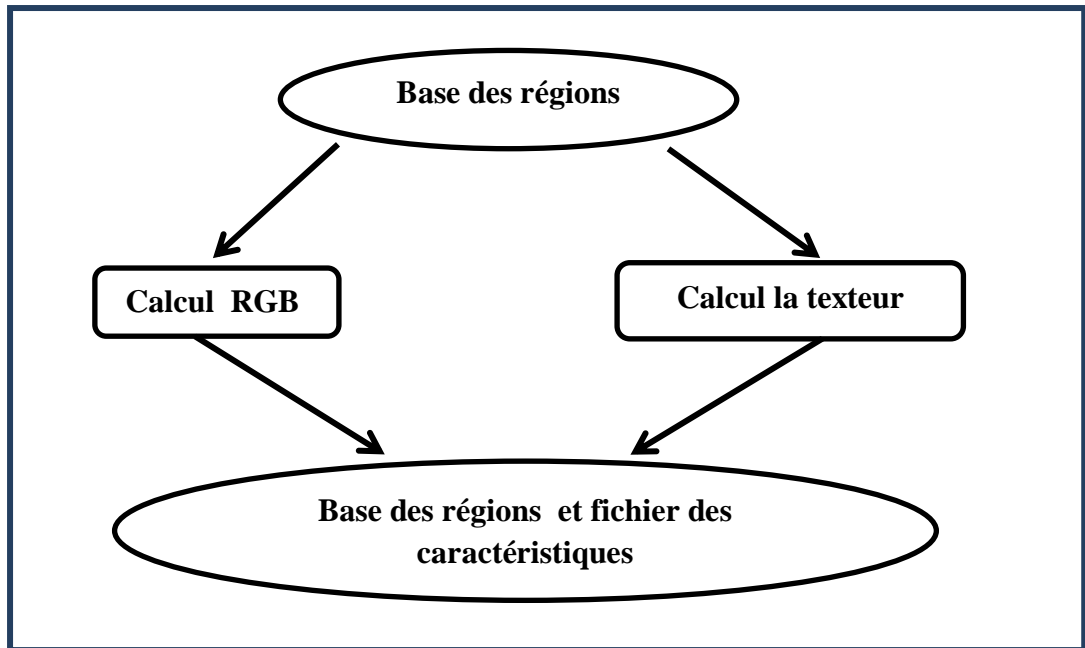


Figure - 21 : Les modules des calculs des caractéristiques des régions.

Ce module se compose de deux sous-modules

#### 3.3.1. Le calcul de la couleur

Nous allons choisir l'histogramme de couleur RGB.

Le calcul de l'histogramme d'une région consiste à parcourir sa matrice de de pixel, et calculer la distribution des couleurs des pixels, c'est-à-dire le nombre de pixels pour chaque valeur de couleur.

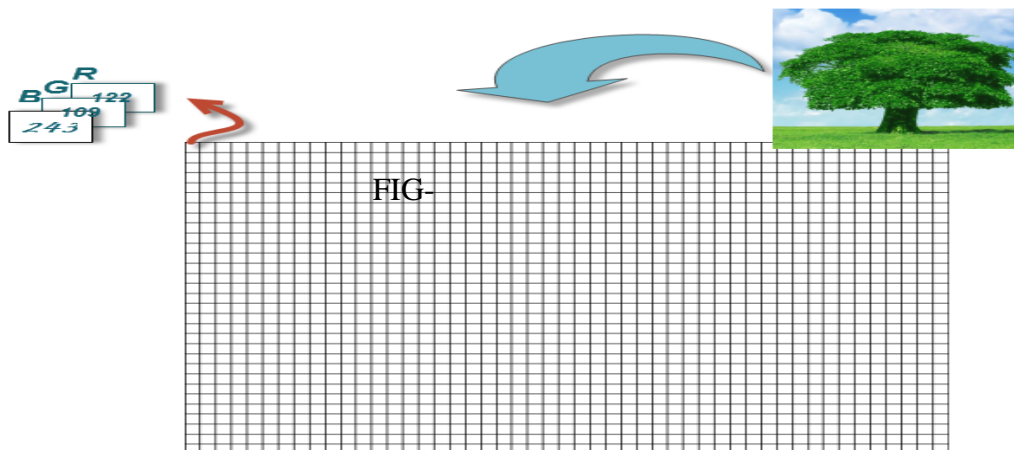


Figure - 22 : calculer de l'histogramme d'une région.

### 3.3.2. Le calcul de la textureur

Le calcul de la texture implique le calcul de la matrice de cooccurrence et les mesures correspondent. Pour cela nous allons passer par trois étapes :

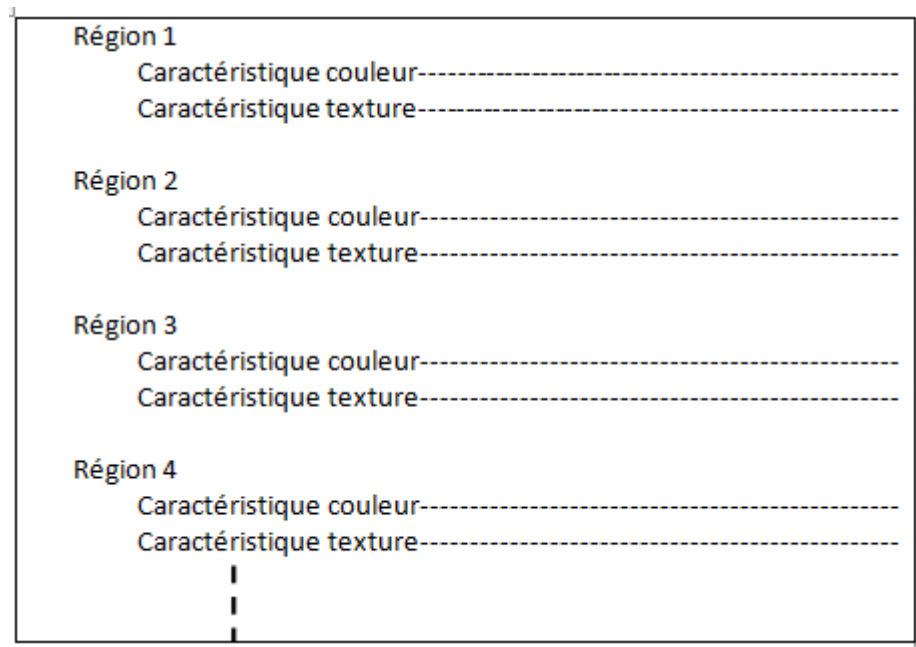
- La conversion de l'image en niveau de gris.
- Le calcul de la matrice de cooccurrence : Le principe de son calcul est le suivant:

MC c'est une matrice d'un volume constante  $255 * 255$  l'opération d'extraction doivent être faite comme suivent:

- Transformation de l'image niveau de Gris
- Sélection de voisinage entre les pixels d'une manière (Horizontal, vertical)
- On prend chaque 2 pixels voisin par rapport les valeurs des deux pixels seraient l'index de colonne et de ligne dans la MC
- Ajoute 1 dans la case qui définir des deux valeurs
- Ont répété cette opération pour chaque fixé plan l'image.
- Le calcul des mesures depuis cette matrice de co-occurrences : comme nous avons présenté dans le deuxième chapitre, il existe 9 mesures. Cependant, les plus utilisées sont : la moyenne, la variance, inertie et la corrélation.

Après le calcul des caractéristiques visuels des régions nous allons les enregistrés dans un fichier texte comme suit :





**Figure - 23 : La structure de fichier de caractéristique.**

### **3.4. Module d'annotation sémantique semi-automatique**

C'est à ce stade que commence le processus d'annotation.

Ce module se compose de deux sous modules:

- Sous module d'annotation manuelle
- Sous module d'annotation automatique

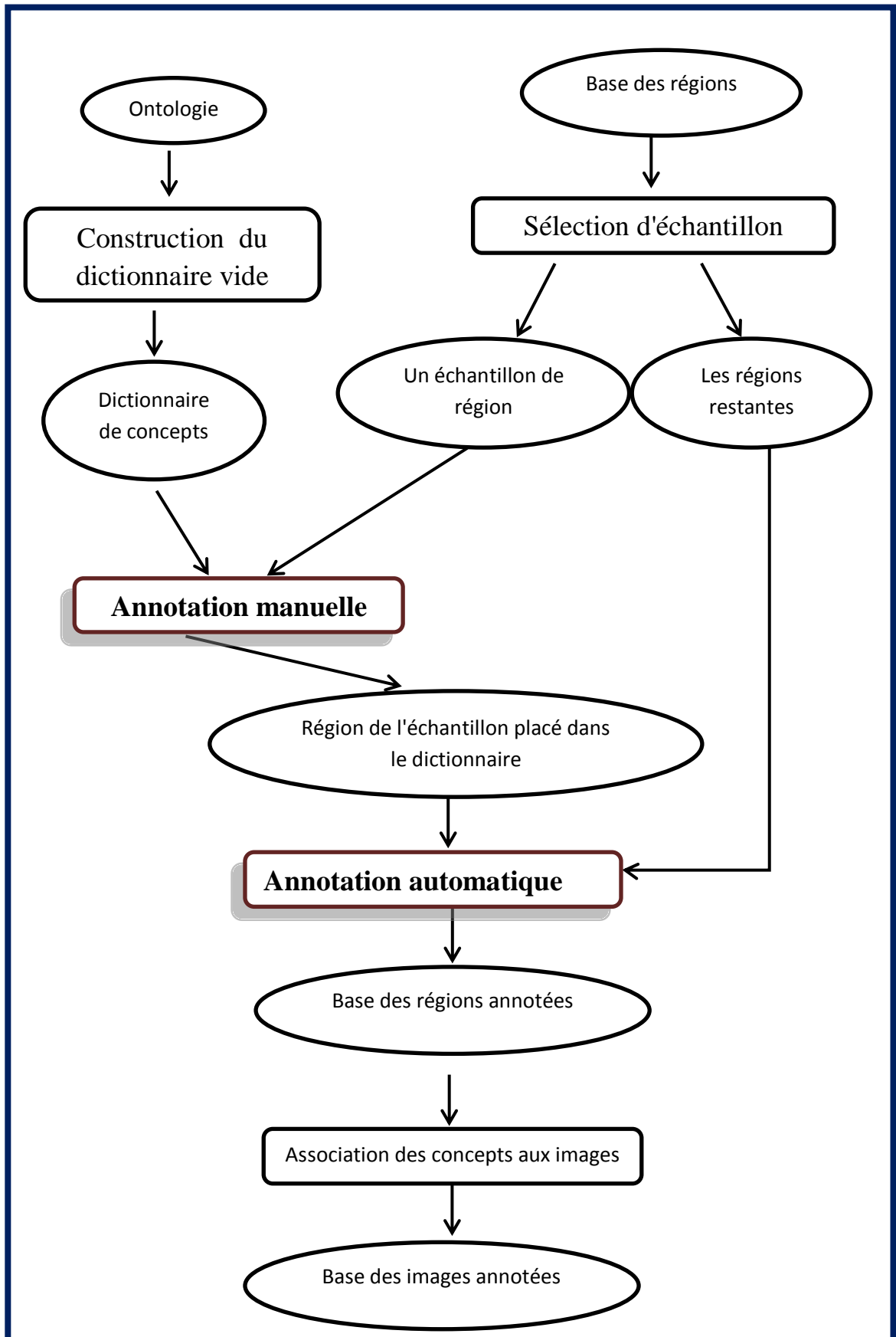


Figure - 24 : Architecture de l'annotation sémantique.

## CHAPITRE IV : Conception

---

Pour commencer, nous allons construire un dictionnaire de concepts de l'ontologie, sous forme d'un répertoire par concept. Initialement il est vide

### 3.4.1. Le sous module d'annotation manuelle

Le but à ce stade est d'avoir des échantillons de régions pour chaque concept du dictionnaire. Pour cela :

- Nous allons sélectionner à notre choix quelques régions de la base des régions.
- Placer manuellement chaque région sélectionnée dans le concept adéquat du dictionnaire.

Il faut signaler que notre sélection des régions échantillons ne serait pas aléatoire mais étudié de tel sorte que :

- Avoir au moins une région par concept
- Choisir des échantillons avec des apparences visuelles différentes pour un même concept.

### 3.4.2. Le sous module d'annotation automatique

Le but de ce module est de placer le reste des régions dans les concepts appropriés d'une façon automatique. Pour cela, nous allons exploiter les caractéristiques visuelles calculées dans le module d'extraction des caractéristiques des régions.

Le principe est de faire une propagation de concepts par mesure de similarité visuelle entre les régions:

- Pour chaque région à annoter, mesurer sa similarité visuelle avec chacune des régions dans le dictionnaire.
- Annoter la région avec le concept pour lequel on obtient une similarité maximale.

Pour comparer les régions nous pouvons exploiter la couleur, la texture, ou bien une combinaison de deux, comme suit :

#### 3.4.2.1. Comparaison de la couleur

Pour comparer la couleur, nous allons exploiter l'histogramme de la couleur RGB. Ainsi, pour mesurer de similarité entre les régions, nous allons utiliser la distance euclidienne.

Une région sera placée dans le concept pour lequel elle a obtenu une distance minimale.

### 3.4.2.2. Comparaison de la texture

La comparaison de la texture des régions consiste à mesurer la similarité des caractéristiques extraites de la matrice de cooccurrence de chaque région non annotée par chacune des régions du dictionnaire.

Comme nous avons présenté auparavant, il existe Neuf Caractéristiques qui peuvent être extraites de la matrice de cooccurrence. Cependant, Les travaux réalisés en ce domaine prouvent que les quatre premières caractéristiques donnent un meilleur résultat.

Comme la couleur, une région sera placée dans le concept pour lequel elle a obtenu une distance minimale.

### 3.4.2.3. Comparaison de la combinaison couleur et texture

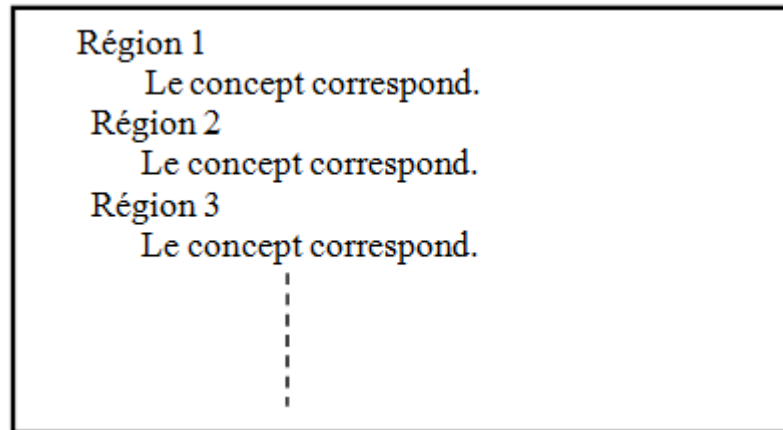
Généralement la combinaison dans le domaine de recherche d'image sert à améliorer le résultat. Pour cela, nous allons l'ajouter pour avoir des mesures plus pertinentes entre les régions.

A ce stade, nous allons exploiter les vecteurs des deux caractéristiques pour chacune des régions. Le principe est le suivant :

- Pour chaque deux région à comparer, nous allons prendre la distance par la couleur et la distance par la texture
- Nous devons trouver une formule pour la combinaison des deux mesures, afin de trouver une distance finale.

A ce niveau, nous serons atteindre le but principal de notre travail, c'est-à-dire que chaque région est placée dans le concept approprié. Autrement dit, elle est annotée par le concept approprié.

Nous allons organiser ce résultat sous forme d'un fichier texte comme suit :



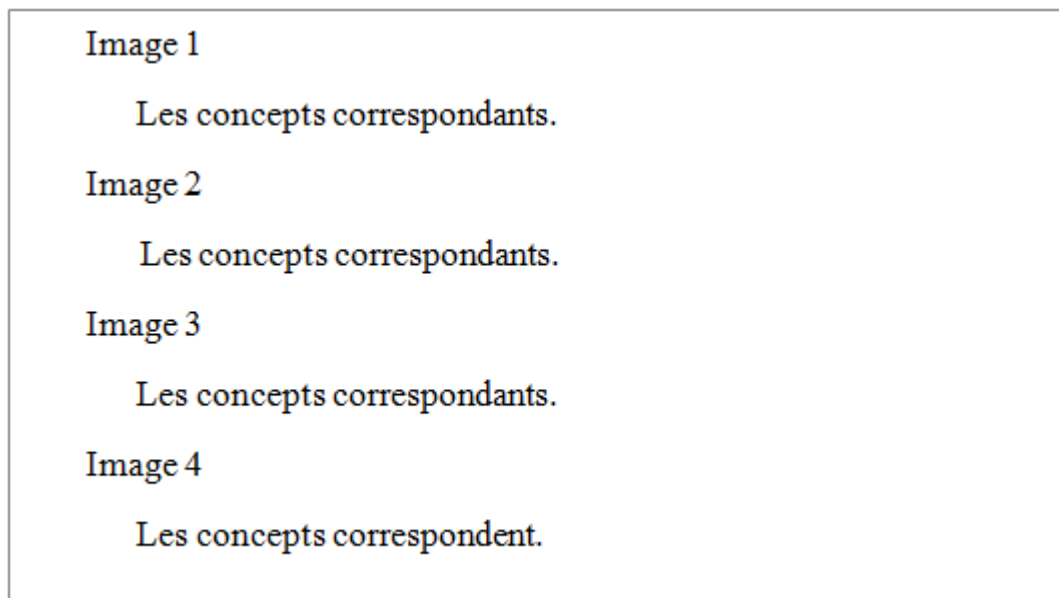
**Figure - 25 : La structure de fichier d'annotation de région.**

Il nous reste de rassembler les annotations pour les images

### **3.4.3. Association des concepts aux images**

Le but de cette étape est de récupérer pour chaque image ses annotations. Le principe est simple :

Pour chaque image, nous allons parcourir la base de régions pour avoir



**Figure - 26 : La structure de fichier d'annotation des images.**

### **4. Conclusion**

Dans ce chapitre, nous avons présenté les étapes détaillées pour la conception de notre système d'annotation sémantique semi-automatique des images. Nous avons commencé par la conception de notre ontologie de domaine *Onto\_Nature*, Ensuite nous avons détaillé les différents modules, le rôle de chacun ainsi que les relations entre ces modules.

Dans le chapitre suivant, nous allons implémenter et mettre en œuvre ce que nous avons proposé dans l'étude conceptuelle. En d'autres termes, la réalisation de notre système d'annotation sémantique semi-automatique des images.

## **CHAPITRE V : Réalisation**

- **Introduction**
- **Construction de notre ontologie**
- **Implémentation de notre système d'annotation  
sémantique semi-automatique**
- **Conclusion**

# CHAPITRE V : Réalisation

## 1. Introduction

Dans ce chapitre nous allons présenter l'implémentation qu'on a fait. Nous présentons alors, dans la première section, le processus de construction de notre ontologie. Par la suite, dans la section deux, nous détaillerons le processus d'implémentation de notre système d'annotation sémantique semi-automatique des images ainsi que des captures d'écrans des différentes interfaces et résultats obtenus.

## 2. Implémentation de notre ontologie « *onto\_nature* » orienté terminologie

### 2.1. Choix de l'éditeur d'ontologie

Pour l'implémentation de notre ontologie, nous avons opté pour l'éditeur d'ontologies (Protégé v3.3.4). Plusieurs raisons ont motivé notre choix :

- C'est un éditeur open source et gratuit.
- Il permet d'importer et d'exporter des ontologies dans les différents langages d'implémentation d'ontologies (RDF-Schéma, OWL, DAML, OIL,...etc.)
- Il possède une interface modulaire, ce qui permet son enrichissement par des modules additionnel (plugins).
- Il permet l'édition et la visualisation d'ontologies.
- Il permet le contrôle de la cohérence de l'ontologie par des vérifications de contraintes.

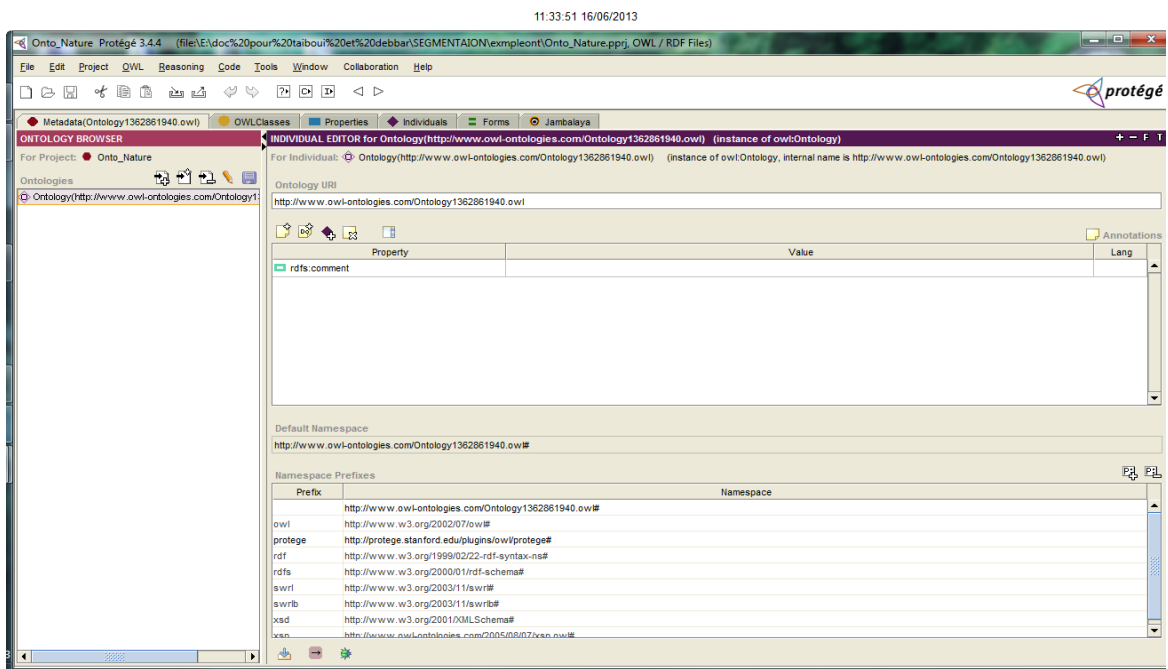


Figure - 27 : Editeur d'ontologie Protégé.



## 2.2. Implémentation de notre ontologie

### 2.2.1. Création des classes

Une fois le projet créé et nommé, on pourrait procéder à la création des classes. Pour ce faire, il faudrait activer l'onglet "Classes". Dans le volet gauche, cliquer sur la racine nommée « THING » avec le bouton droit de la souris. Ce qui aura pour effet d'afficher un menu contextuel. L'utilisateur devrait alors cliquer sur la commande "Creat Subclass", qui affichera une zone de texte dans laquelle, il saisira le nom de la classe. Celle-ci se place directement sous la racine.

Voici une capture d'écran des classes de notre ontologie

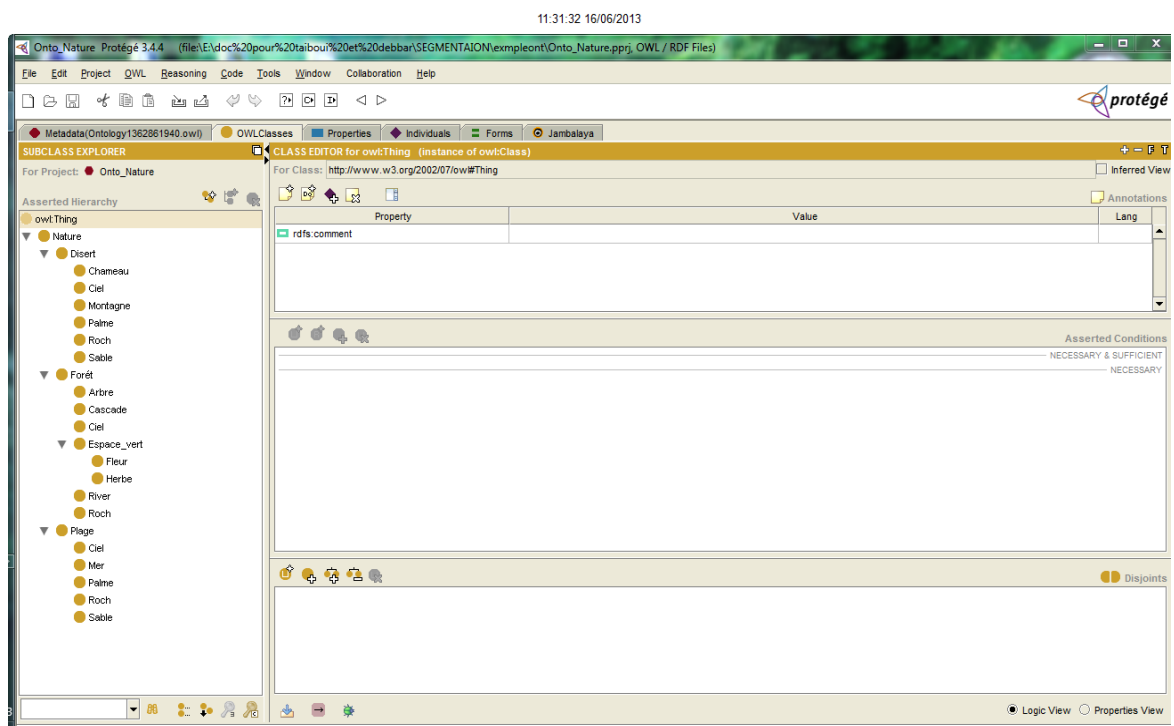


Figure - 28 : Description des Classes d'onto\_nature.

### 2.2.2. Création des relations

La création des relations se fait en sélectionnant la classe dans le volet gauche de l'arborescence, puis en cliquant sur l'icône situé dans la zone " ObjectProperties", dans le volet droit. Une boîte de dialogue s'affiche, l'utilisateur est invité alors de saisir les valeurs des différentes zones de texte.

## CHAPITRE V : Réalisation

Dans notre ontologie nous avons définies des relations de type *partie\_de* entre les différents concepts

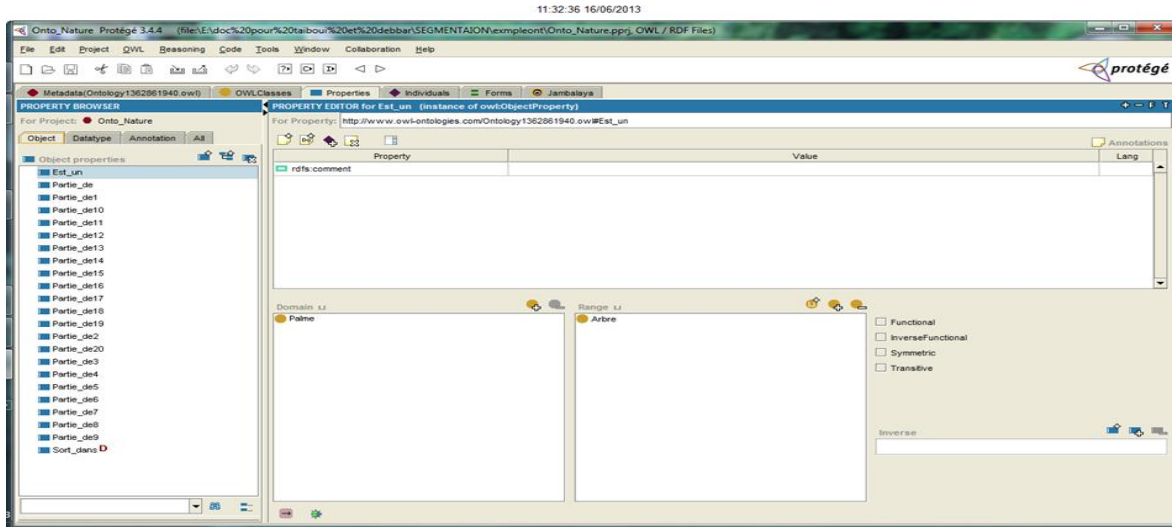


Figure - 29 : Création des relations.

### 2.2.3. Visualisation de l'ontologie

Protégé offre la possibilité de visualiser graphiquement une ontologie.

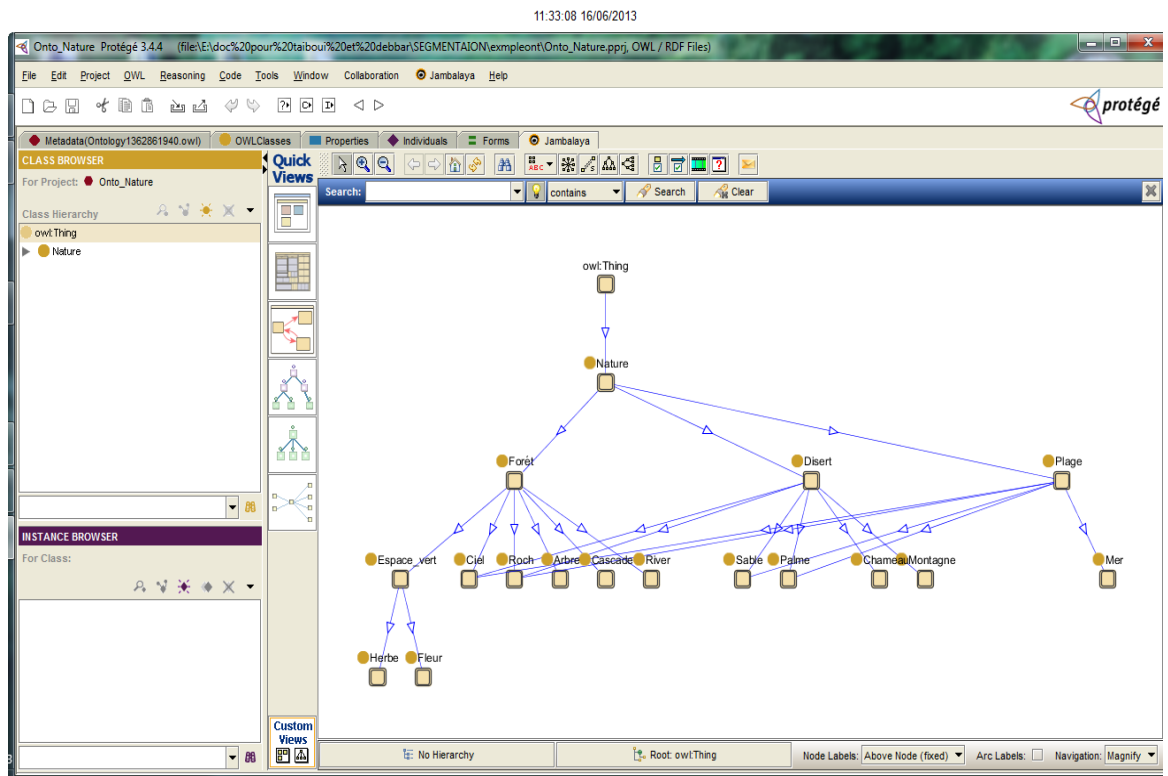
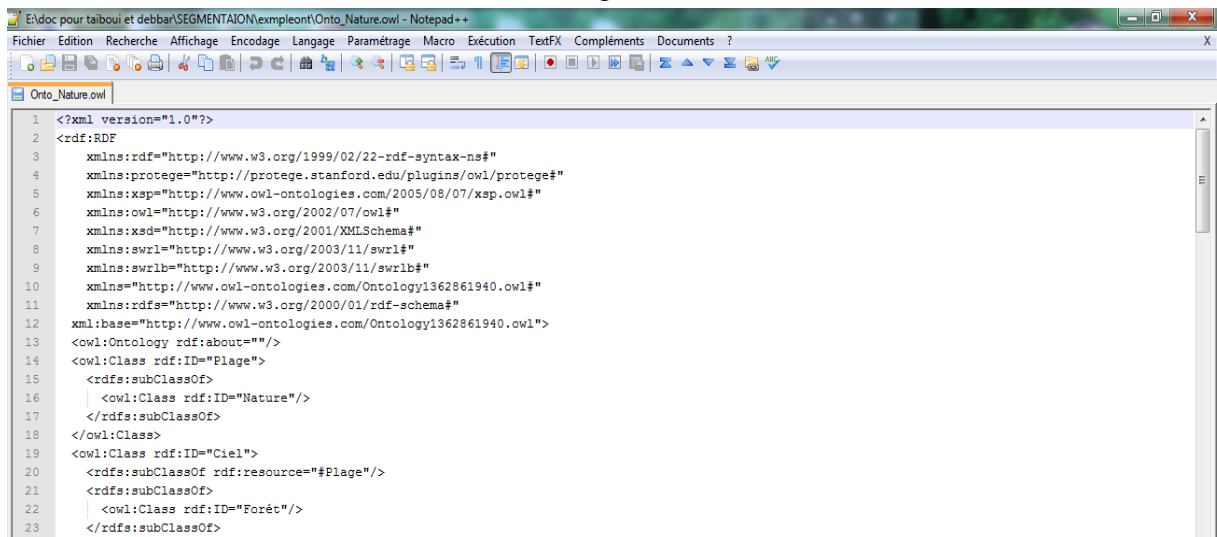


Figure - 30 : Visualisation de l'ontologie par Protégé.

Voici un extrait du code OWL de notre ontologie



```
1 <?xml version="1.0"?>
2 <rdf:RDF
3   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
4   xmlns:protege="http://protege.stanford.edu/plugins/owl/protege#"
5   xmlns:xsp="http://www.owl-ontologies.com/2005/08/07/xsp.owl#"
6   xmlns:owl="http://www.w3.org/2002/07/owl#"
7   xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
8   xmlns:swrl="http://www.w3.org/2003/11/swrl#"
9   xmlns:swrlb="http://www.w3.org/2003/11/swrlb#"
10  xmlns="http://www.owl-ontologies.com/Ontology1362861940.owl#"
11  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
12  xml:base="http://www.owl-ontologies.com/Ontology1362861940.owl">
13 <owl:Ontology rdf:about=""/>
14 <owl:Class rdf:ID="Flage">
15   <rdfs:subClassOf>
16     <owl:Class rdf:ID="Nature"/>
17   </rdfs:subClassOf>
18 </owl:Class>
19 <owl:Class rdf:ID="Ciel">
20   <rdfs:subClassOf rdf:resource="#Flage"/>
21   <rdfs:subClassOf>
22     <owl:Class rdf:ID="Forêt"/>
23   </rdfs:subClassOf>
```

Figure - 31 : le code OWL de notre ontologie « Onto\_Nature ».

### 3. L'implémentation de notre système d'annotation sémantique semi-automatique des images

Avant de commencer l'implémentation de notre système, nous allons présenter rapidement les différents outils que nous avons utilisés.

#### 3.1. Choix des outils

##### 3.1.1. Visual Studio

Microsoft Visual Studio est une suite de logiciels de développement pour Windows conçue par Microsoft. La dernière version s'appelle Visual Studio 2012. Visual Studio est un ensemble complet d'outils de développement permettant de générer des applications Web ASP.NET, des Services Web XML, des applications bureautiques et des applications mobiles. Visual Basic, Visual C++, Visual C# et Visual J# utilisent tous le même environnement de développement intégré (IDE, Integrated Development Environment), qui leur permet de partager des outils et facilite la création de solutions faisant appel à plusieurs langages. Par ailleurs, ces langages permettent de mieux tirer parti des fonctionnalités du Framework .NET, qui fournit un accès à des technologies clés simplifiant le développement d'applications Web ASP et de Services Web XML grâce à Visual Web Developer.

Pour la programmation de notre système nous avons opté pour le langage C# :

## CHAPITRE V : Réalisation

---

### 3.1.2. C Sharp (c#)

C sharp (c#) est un langage récent. Il a été disponible en versions beta successives depuis l'année 2000 avant d'être officiellement disponible en février 2002 en même temps que la plate-forme .NET 1.0 de Microsoft à laquelle il est lié. C# ne peut fonctionner qu'avec cet environnement d'exécution. Celui-ci rend disponible aux programmes qui s'exécutent en son sein un ensemble très important de classes. En première approximation, on peut dire que la plate-forme .NET est un environnement d'exécution analogue à une machine virtuelle Java.

### 3.1.3. DevExpress

Nous avons utilisé DevExpress pour profiter de ses composants graphiques de bonne qualité. Il fournit gracieusement 40 de ses composants (XtraEditors Library, XtraNavBar, ASPxMenu and ASPxSiteMapControl), dans le cadre d'une utilisation non commerciale.

### 3.1.4. Adobe Photoshop

Logiciel de retouche d'images développé et fabriqué par Adobe System Inc. Photoshop est considéré comme l'un des leaders dans les logiciels de retouche photo. Le logiciel permet aux utilisateurs la manipulation, le recadrage, le redimensionnement,...etc des photos numériques.

## 3.2. Les étapes d'implémentation

### 3.2.1. Le choix d'une base d'images

Nous avons créé une base d'images, sous forme d'un répertoire, de 54 images de la nature collectées depuis l'Internet. Les images sont sélectionnées en fonction des concepts de notre ontologie.

### 3.2.2. La segmentation de la base d'images

Comme nous avons mentionné dans le chapitre précédent, on va commencer par l'implémentation de filtrage.

#### 3.2.2.1. L'implémentation de filtrage d'images

FI (filtrage d'image) est une petite application qui permet de filtrer les images selon le principe décrit précédemment.

Elle contient les étapes suivantes :

## CHAPITRE V : Réalisation

1. Sélection d'une base d'images par un clic sur le bouton «choisir les images ».
2. Saisir la valeur du filtrage.
3. Lancement de filtrage et choix un emplacement pour les images filtrées, en clique sur le bouton « filtrer ».

La figure suivante montre les étapes précédentes :



Figure - 32 :L'interface de filtrage.

### 3.2.2.2. Implémentation de la segmentation des images filtrées

La segmentation s'applique pour chaque image de la base de données. Elle consiste à extraire les régions de l'image. On obtient comme résultat une base de toutes les régions des images.

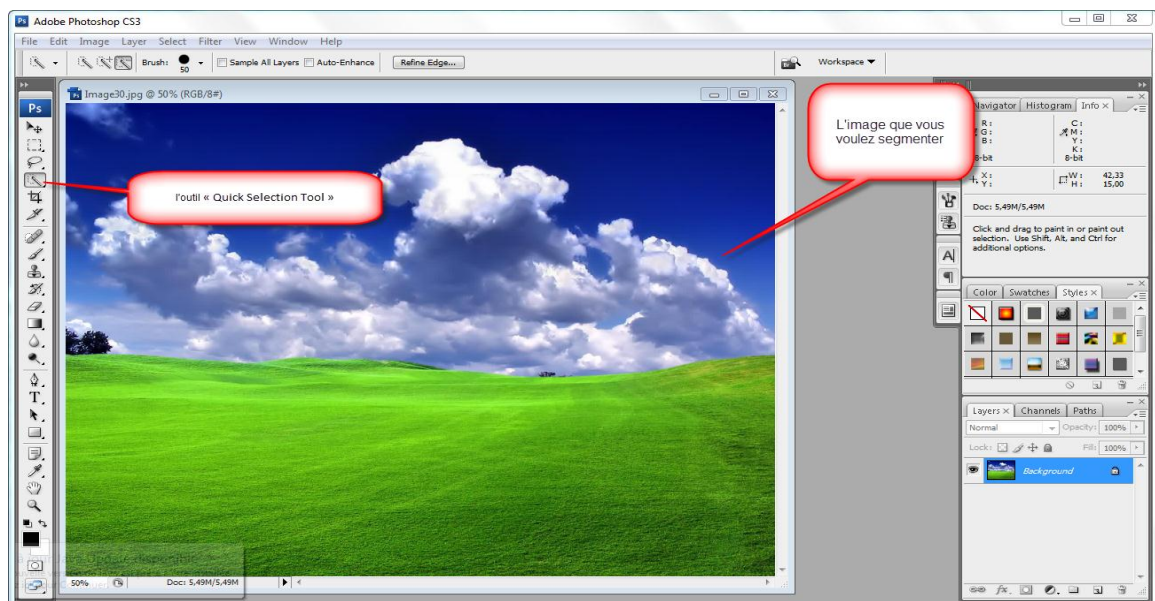


Figure - 33 : Sélection des régions via l'outil « Quick Selection Tool ».

## CHAPITRE V : Réalisation

Nous avons procédé la segmentation d'une façon manuelle par l'utilisation du logiciel *Adobe Photoshop*.

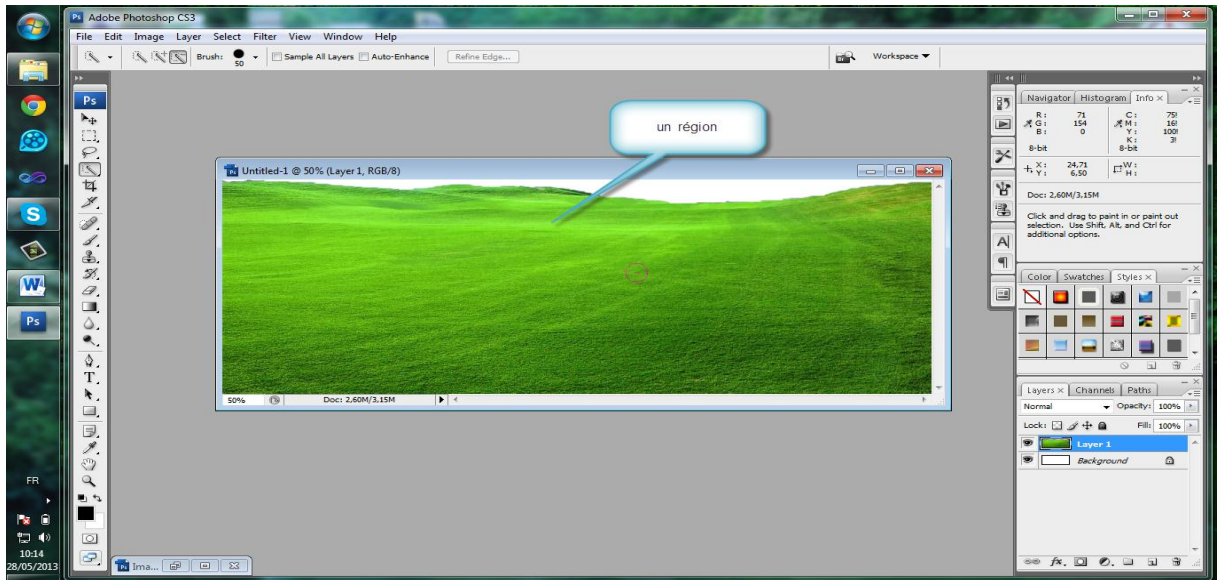


Figure - 34 : Une région avec un arrière-plan blanc.

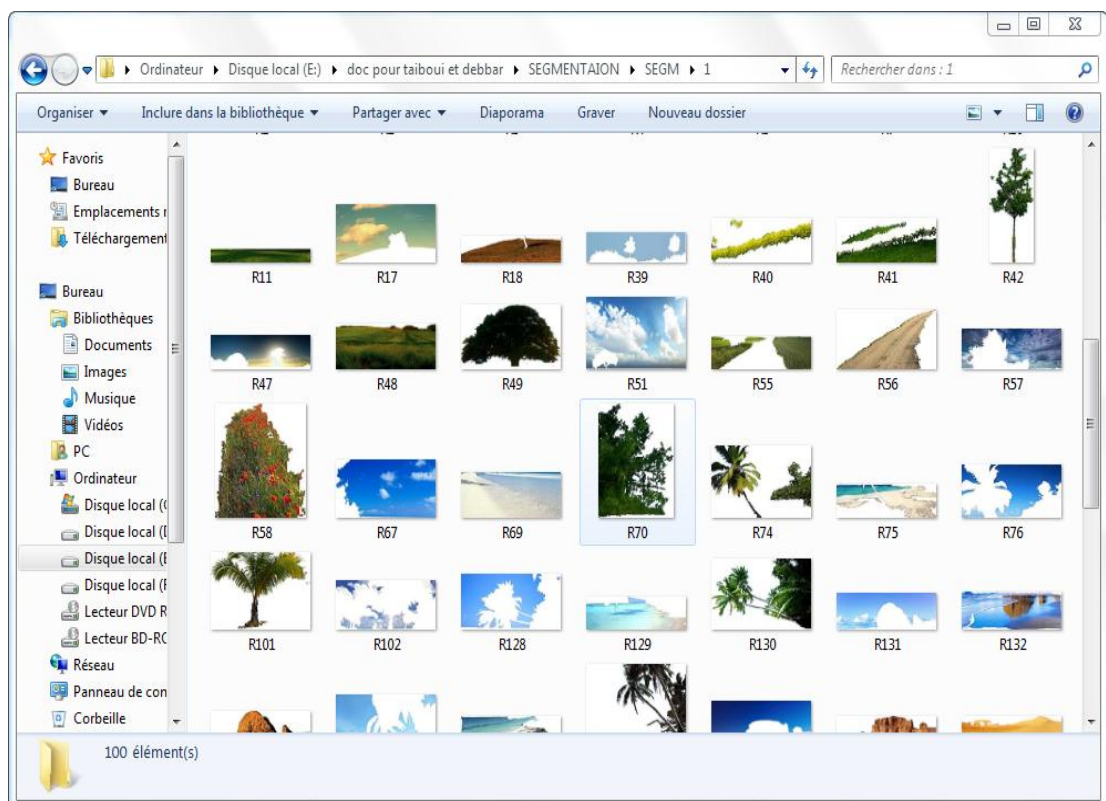


Figure - 35 : Le résultat de segmentation, la base de régions.

## CHAPITRE V : Réalisation

A ce stade, nous avons le nécessaire pour commencer notre travail proprement dit sur l'annotation semi-automatique.

Notre application s'appelle ASSA (**a**nnotation **s**émantique **s**emi-**a**utomatique). On peut la lancer grâce à l'interface suivante:

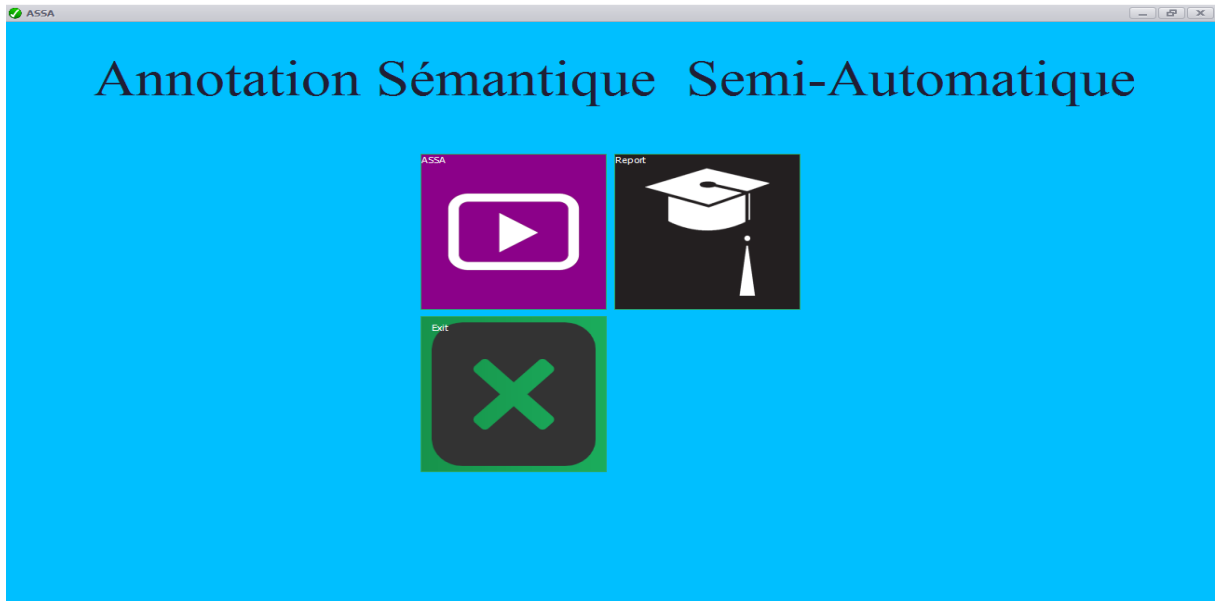


Figure - 36 : Interface pour lancer l'application.

### 3.2.3. L'interface principale de notre application

L'interface illustrée par la figure ci-dessous représente l'interface principale de notre système ASSA d'annotation sémantique semi-automatique des images.

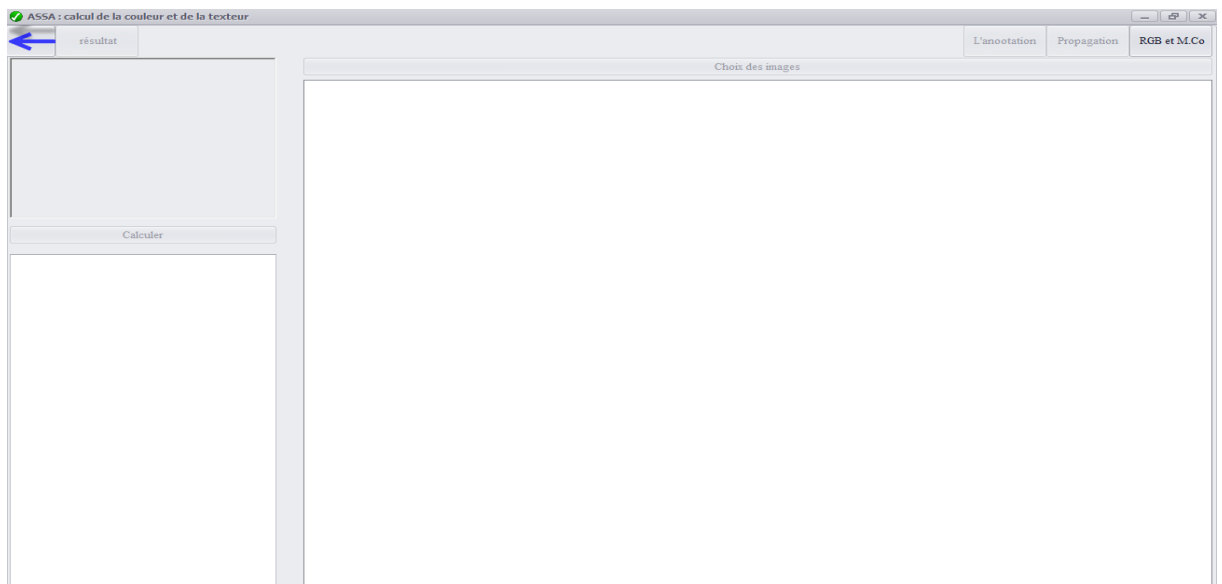


Figure - 37 : l'interface principale.

## CHAPITRE V : Réalisation

---

### 3.2.4. Le calcul des caractéristiques visuelles des régions

Nous avons implémenté le calcul de deux caractéristiques visuelles très utilisés : la couleur et la texture. De plus nous avons implémenté leur combinaison.

Ce travail était incrémental comme suit :

#### 3.2.4.1. Calcul de l'histogramme RGB de chaque région

Pour calculer l'histogramme d'une région il faut parcourir son matrice de pixel et calculer le nombre de pixels pour chaque valeur de couleur. Avec la connaissance que chaque pixel est composé de trois valeurs (R, G, B). Il ne faut pas oublier qu'il faut ignorer le fond de la région pendant ce calcul (le but de filtrage)

Donc, partant d'une matrice de pixels, nous appliquons l'algorithme suivant :

```
ALGORITHME calcul RGB
  Tableaux RGB [27] :intègre ;
  Matrice[n,m] MathPixel :intègre ;
  R,G,B,i,j ,Position :intègre ;
Début
  Lire(image) ;
  MathPixel = transformer_Matrice_Pixel(image) ;
  Pour i = 1 to MathPixel.Large faire
    Pour j =1 to MathPixel.long faire
      Si (MathPixel[i,j].R != 255 ET MathPixel[i,j].G != 255 ET
          MathPixel[i,j].B !=255 )
        R = MathPixel[i,j].R / 86 ;
        G = MathPixel[i,j].G / 86 ;
        B = MathPixel[i,j].B / 86 ;
        Position = ( ( R * 9 ) + ( G * 3 ) + B ) ;
        RGB[ Position] = RGB[ Position]+1 ;
      Finsi
    Fin Pour
  Fin Pour
  Ecrire RGB ;
Fin.
```



## CHAPITRE V : Réalisation

---

### 3.2.4.2. Le Calcul de la matrice de cooccurrence pour chaque région

Pour calculer la matrice de cooccurrence et les mesures qu'on peut extraire de cette matrice, nous avons passé par trois étapes :

- La conversion de l'image en niveau de gris.

```
Algorithme      niveau de gris

  l,j ,val: entier ;

  Img : type image ;

Debut

  Lire (img) ;

  Pour i=1 a img.longueur faire

    Pour j=1 a img.larguer faire

      Val =(img[i,j].R+ img[i,j].G+ img[i,j].B )/3;

      img[i,j]= val ;

    FInPour

  FInPour

  Return img

Fin
```

- Le calcul la matrice de cooccurrence : Partant d'une matrice de pixels, nous appliquons l'algorithme suivant :

## CHAPITRE V : Réalisation

---

ALGORITHME calcul Matrice cooccurrence

Matrice[256,256] MathCooc :entier ;

Matrice[n,m] MathPixel :entier ;

X , Y , i,j ,Position :entier ;

Début

Lire(image) ;

MathPixel = convertir\_niveau\_degris (transformer\_Matrice\_Pixel(image)) ;

Pour i = 1 to MathPixel.Large faire

    Pour j =1 to ( MathPixel .long -1) faire

        Si (MathPixel[i , j] != 255 ET MathPixel[i , j] != 255 )

            X= MathPixel[i , j] ;

            Y= MathPixel[i , j+1] ;

            MathCooc [ X,Y ]= MathCooc [X,Y] + 1 ;

        Fisi

    Fin Pour

Fin Pour

Ecrire MathCooc ;

Fin

Le calcul des mesures depuis cette matrice de cooccurrence : Nous avons calculé les mesures les plus significatives à savoir la moyenne, la variance, l'inertie et la corrélation

Dans ce qui suit, C'est la matrice de cooccurrence, **i** est l'indice des lignes dans cette matrice et **j** l'indice des colonnes.  $\mu$  est la moyenne (c'est-à-dire la première caractéristique *Mean*).

$$Mean = \sum_i \sum_j i.C(i, j)$$

$$Variance = \sum_i \sum_j (i - \mu)^2.C(i, j)$$

$$Energy = \sum_i \sum_j C(i, j)^2$$

## CHAPITRE V : Réalisation

$$Correlation = \sum_i \sum_j (i - \mu)(j - \mu).C(i, j)$$

Après le calcul des caractéristiques visuels des régions nous avons les enregistrés dans un fichier texte

Dans la figure ci-après une capture de l'interface de calcul de ces caractéristiques

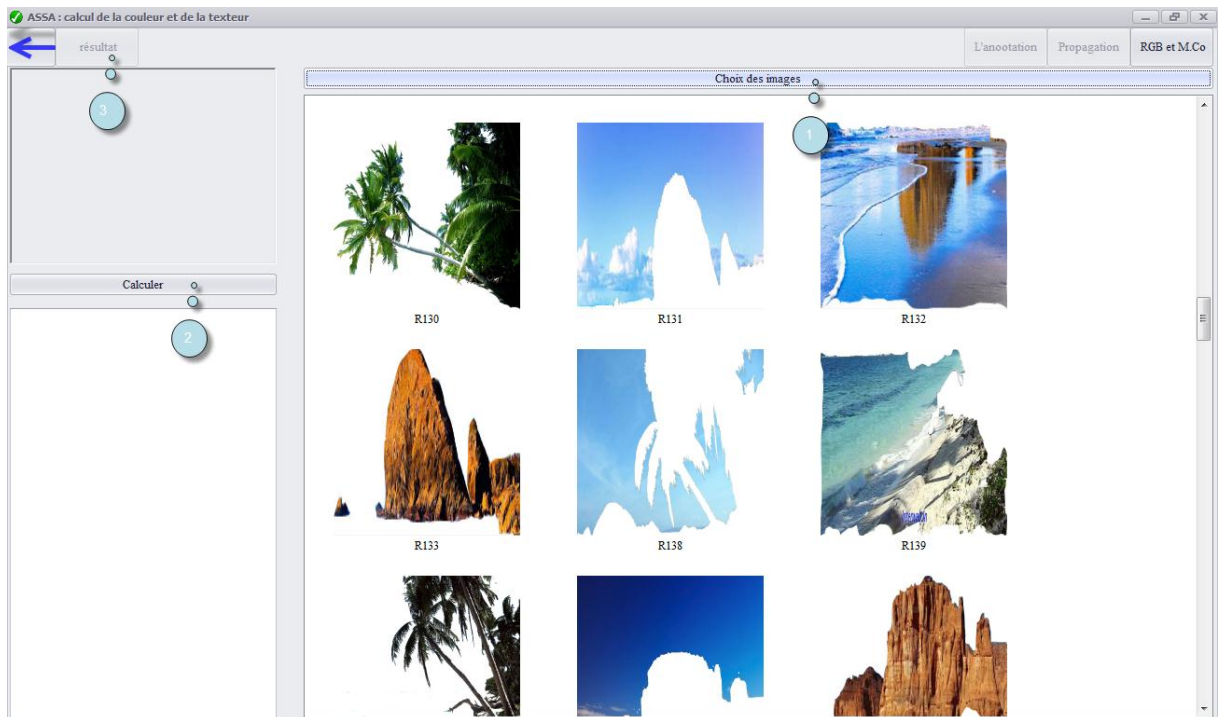


Figure - 38 : l'interface de calcul des caractéristiques des régions.

Pour procéder :

1. Sélectionner la base de régions.
2. Lancer le calcul des caractéristiques, avec spécification du nom de fichier résultat qui va contenir ces caractéristiques.
3. Affichage de fichier des caractéristiques.

L'image suivante présenter un exemple sur le fichier de caractéristiques :

## CHAPITRE V : Réalisation

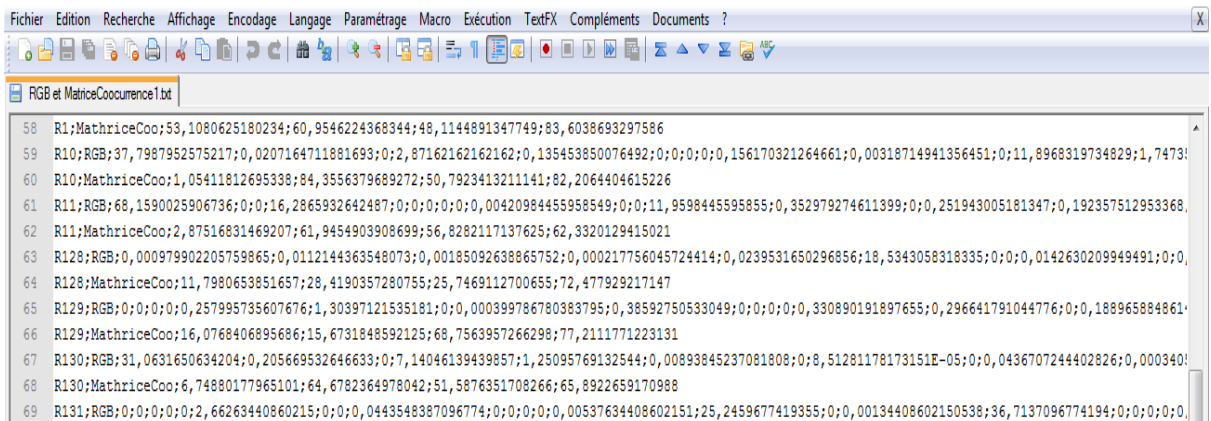


Figure - 39 : un exemple de fichier des caractéristiques.

### 3.2.5. Implémentation de l'annotation sémantique semi-automatique

L'implémentation de l'annotation semi-automatique à base d'ontologie comprend les étapes suivantes :

#### 3.2.5.1. La création d'un dictionnaire

Nous avons créé un dictionnaire de concepts de notre ontologie, sous forme d'un répertoire par concept. Initialement il est vide.

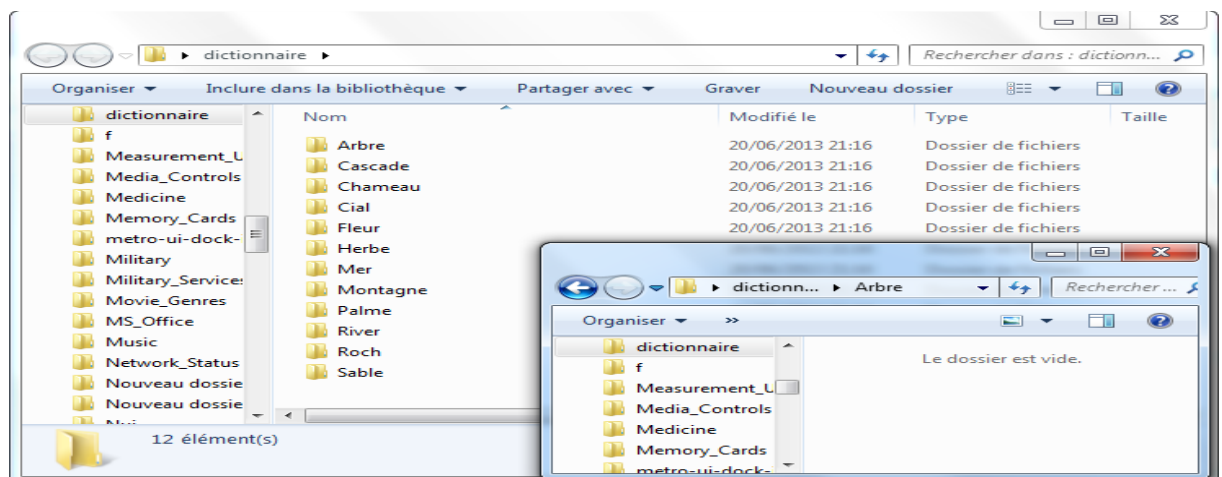


Figure - 40 : dictionnaire vide.

#### 3.2.5.2. L'annotation manuelle d'un échantillon de régions

Pour placer un échantillon de régions pour chaque concept dans le dictionnaire, nous avons :

## CHAPITRE V : Réalisation

- Sélectionner à notre choix quelques régions de la base des régions, de tel sorte :
  - o Avoir au moins une région par concept
  - o Choisir des échantillons avec des apparences visuelles différentes pour un même concept.
- Placer manuellement chaque région sélectionnée dans le concept adéquat du dictionnaire.

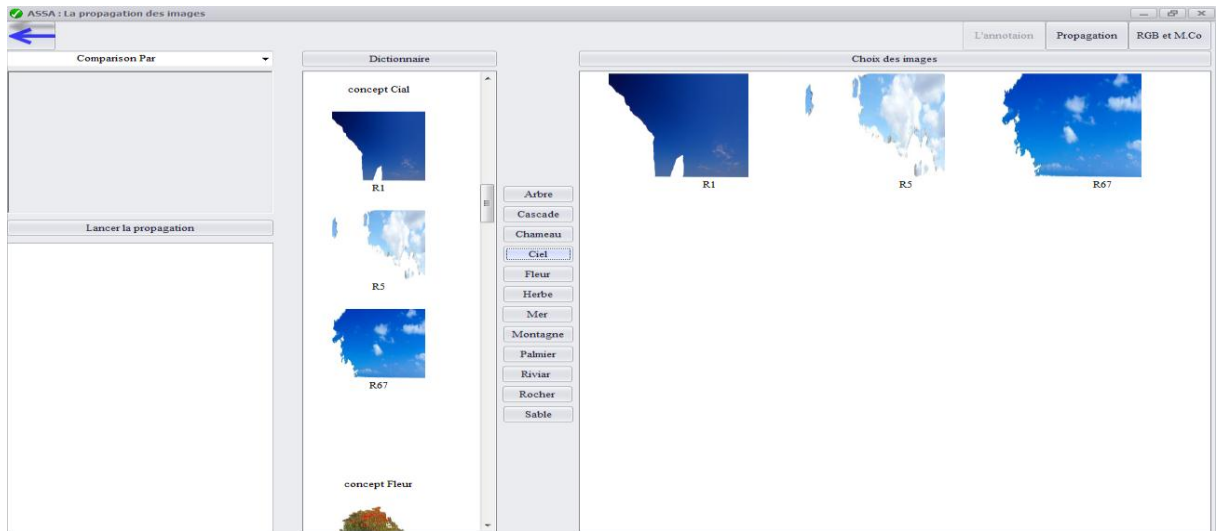


Figure - 41 : Un échantillon de régions dans le concept ciel du dictionnaire.

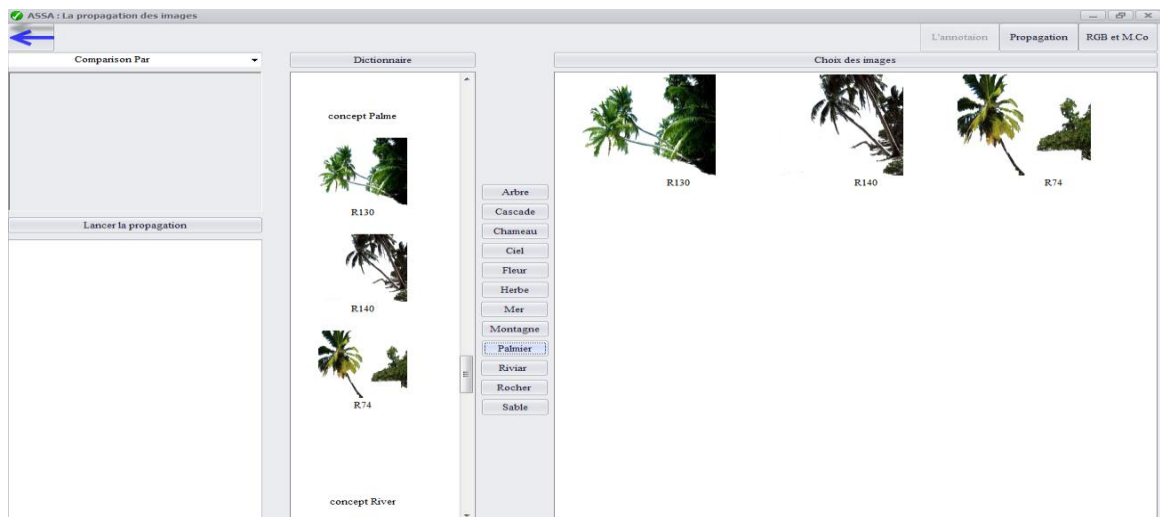


Figure - 42 : Un échantillon de régions dans le concept palmier du dictionnaire.

## CHAPITRE V : Réalisation

### 3.2.5.3. L'annotation automatique des régions restantes

Pour placer le reste des régions dans les concepts appropriés d'une façon automatique, nous avons procédé une propagation de concepts par mesure de similarité visuelle entre les régions, en exploitants les caractéristiques visuelles calculées précédemment:

- Pour chaque région à annoter, mesurer sa similarité visuelle avec chacune des régions dans le dictionnaire.
- Placer la région dans le répertoire du concept pour lequel on obtient une similarité maximale.

Pour ce faire, l'utilisateur doit :

1. Sélectionner la base de régions.
2. Importer le dictionnaire.
3. Choisir les caractéristiques visuelles à utiliser pour faire la propagation.
4. Lancer la propagation.

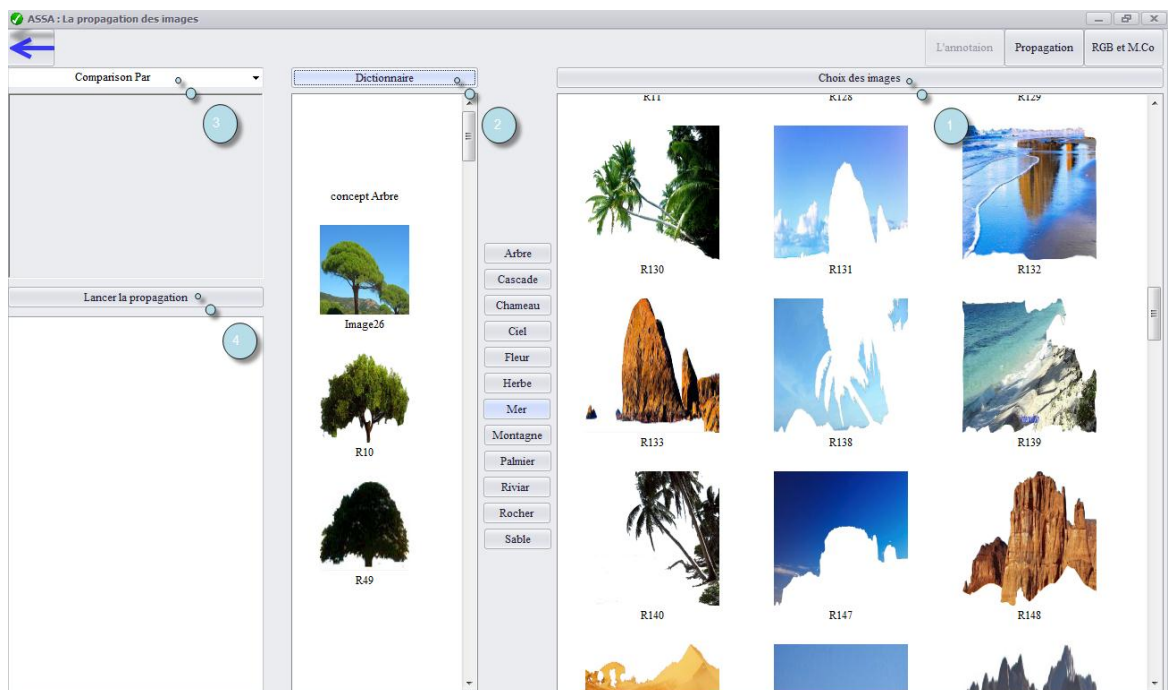


Figure - 43 : Dictionnaire avant la propagation des concepts.

L'exploitation des caractéristiques visuelles pour faire la propagation des concepts était incrémentale comme suit:

## CHAPITRE V : Réalisation

---

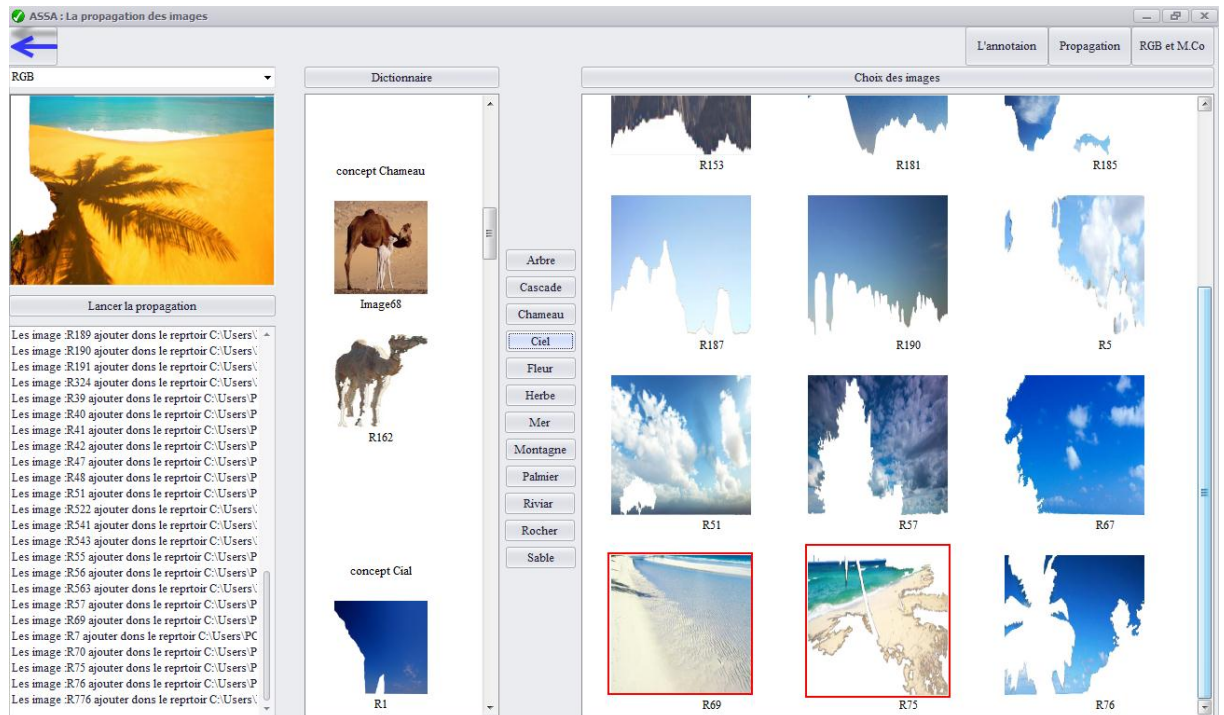
### A. Utilisation de la couleur

Nous avons défini précédemment le principe de cette comparaison, à partir de fichier des caractéristiques, nous appliquons l'algorithme suivant :

```
Algorithme comparaison par couleur
Tableaux V_RGB1, V_RGB2 : entier ;
Min_RGB, voleur_RGB : entier ;
List Résulte_D_C :entier ;
Début
Lire (fiche de caractéristiques) ;
Lire (image)
//vecteur RGB est un fonction qui return vecteur RGB de image en fichier
caractéristique
V_RGB1 = vecteur_RGB (En_image)
Pour I = 1 à dictionnaire. Longueur faire
V_RGB2 = vecteur_RGB(dictionaries [i])
// Distance_ecludian est un fonction calcul la Distance ecludian entre
V_RGB1 et V_RGB2
valeur_RGB = Distance_ecludian (V_RGB1,V_RGB2);
Result_D_C ajout(valeur_RGB)
fin pour
//fonction qui return le minimum valeur dans list
Min_RGB = Min (Result_D_C)
// je vi ajouter l'image En_image dans le dictionaries
Diction.Ajout(image[Min_RGB])
Fin
```

## CHAPITRE V : Réalisation

Les figures suivantes montrent des résultats obtenus :



**Figure - 44: Résultat Après mesure de similarité visuelle de la couleur.**

Nous avons constaté que certaines régions sont placées dans les concepts adéquats. Cependant, d'autres concepts sont mal placés, par exemple des images de la mer sont placées avec des images du ciel. Ceci est évident car leurs couleurs sont similaires.

### Critique

L'utilisation de la couleur seule à provoquer pour, certaines régions, une fausse classification dans le dictionnaire des concepts. Donc, elle est insuffisante pour placer les régions dans les concepts adéquats du dictionnaire. Pour cela, nous avons exploité une autre caractéristique, c'est la texture.

### **B. Comparaison de la texture**

Après le calcul des quatre mesures : la moyenne, la variance, l'inertie et la corrélation, nous avons calculé les distances entre ces mesures. La distance utilisée est la distance euclidienne.



## CHAPITRE V : Réalisation

---

```
Algorithme comparaison par textureur
Tableaux V_text1, V_text2 : entier ;
valeur_text , Min_text  : entier ;
List Resulte_D_C :entier ;
Debut
    Lire (fiche de caractéristiques) ;
    Lire (text_image)
    //vecteur_texteur est un fonction qui return vecteur texture de image en fichier
    caractéristique
    V_text1 = vecteur_texture (text_image)
    Pour I = 1 à dictionnaire. Longueur faire
        V_text = vecteur_texture (dictionnaire[i])
        // Distance _ecludian est un fonction calcul la Distance ecludian entre
        V_text1, V_text2
        valeur_text = Distance _cludian (V_text1, V_text2 );
        Result_D_C ajout(valeur_text)
    fin pour
    //fonction qui return le muninent valeur dans list
    Min_text = Min (Result_D_C)
    // je vi ajouter l'image text_image dans le dictionaries
    Diction.Ajout(imge[Min_text])
Fin
```

# CHAPITRE V : Réalisation

Les figures suivantes montrent les résultats obtenus

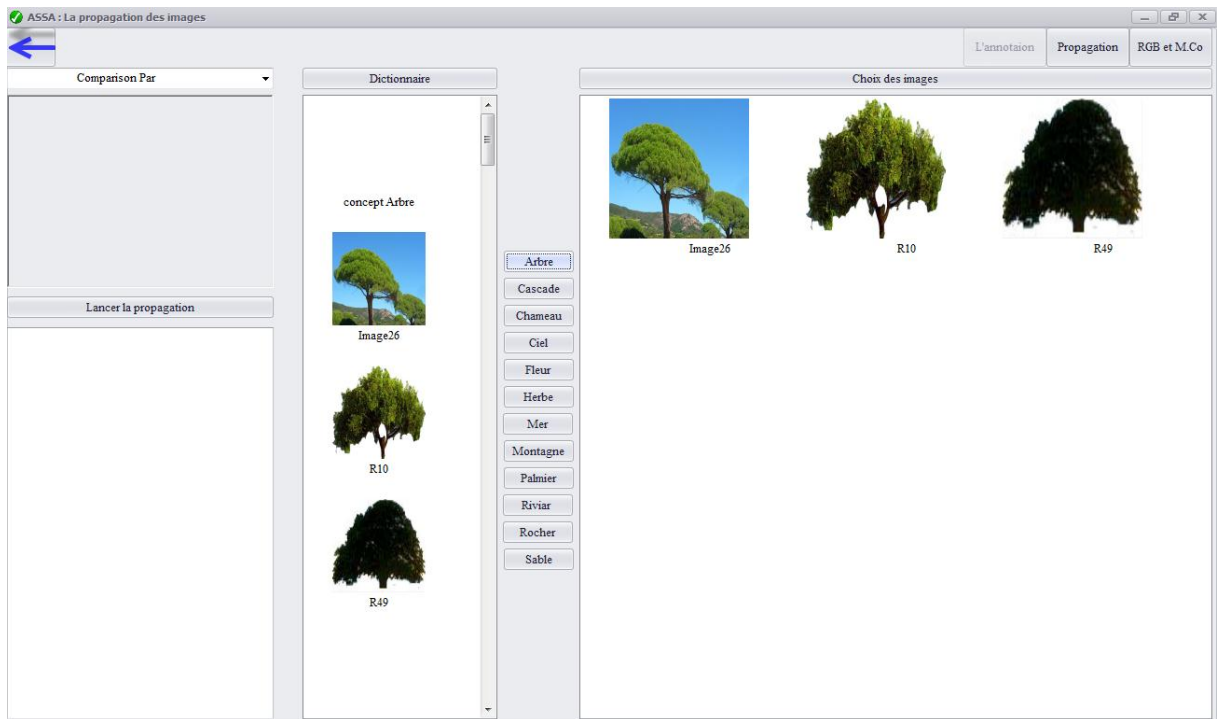


Figure - 45 : Le contenu du concept arbre avant.

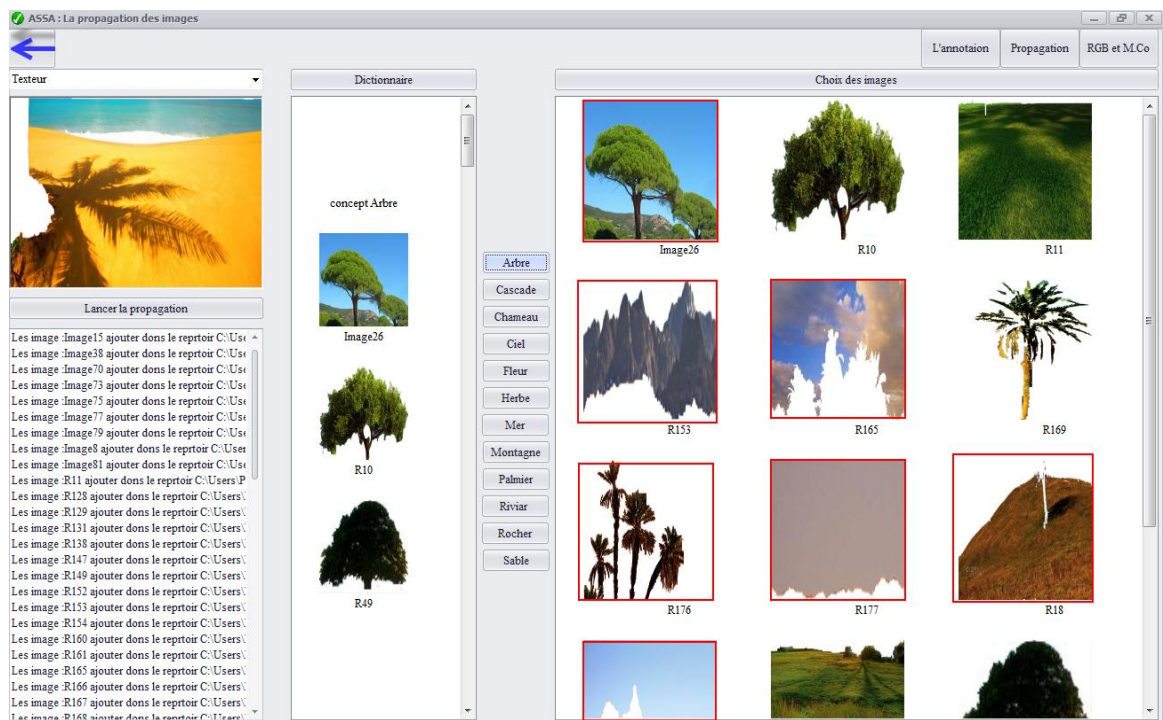


Figure - 46 : Résultat Après mesure de similarité visuelle de la texture

## CHAPITRE V : Réalisation

### Critique

L'utilisation de la texture seule donne un résultat médiocre. Donc, elle est insuffisante pour placer les régions dans les concepts adéquats du dictionnaire.

A cet effet, nous avons trouvé qu'il est indispensable de faire une combinaison des deux caractéristiques couleurs et texture

### C. Comparaison de la combinaison couleur et texture

Pour mesurer la similarité de deux régions, en fonction des de ces deux caractéristiques, nous avons :

- Calculer la distance euclidienne pour la couleur ( $d_1$ )
- Calculer la distance euclidienne pour la texture ( $d_2$ )
- Calculer une distance finale par la formule :  $\text{poid}_i * d_1 + \text{poid}_j * d_2$

Les figures suivantes montrent des exemples des résultats obtenus :

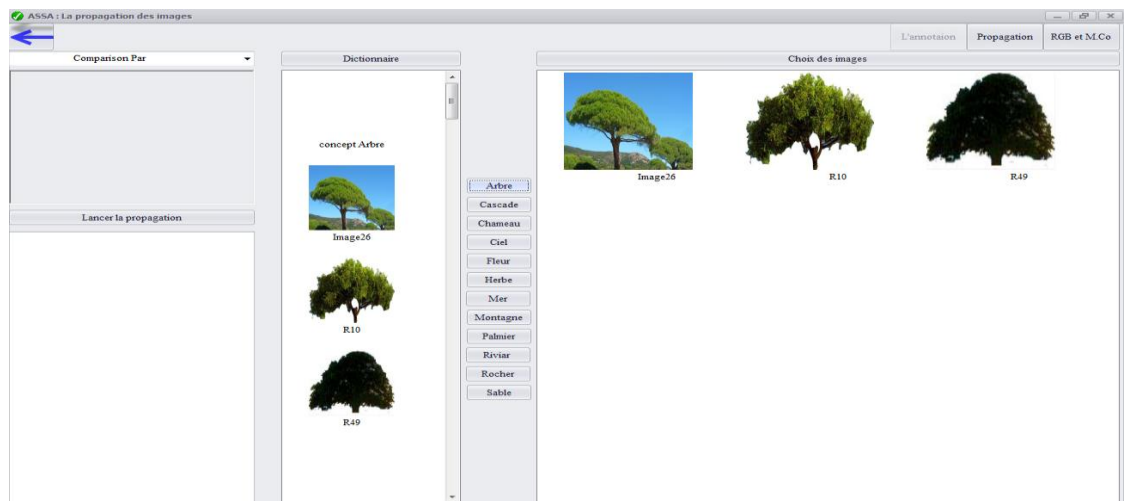
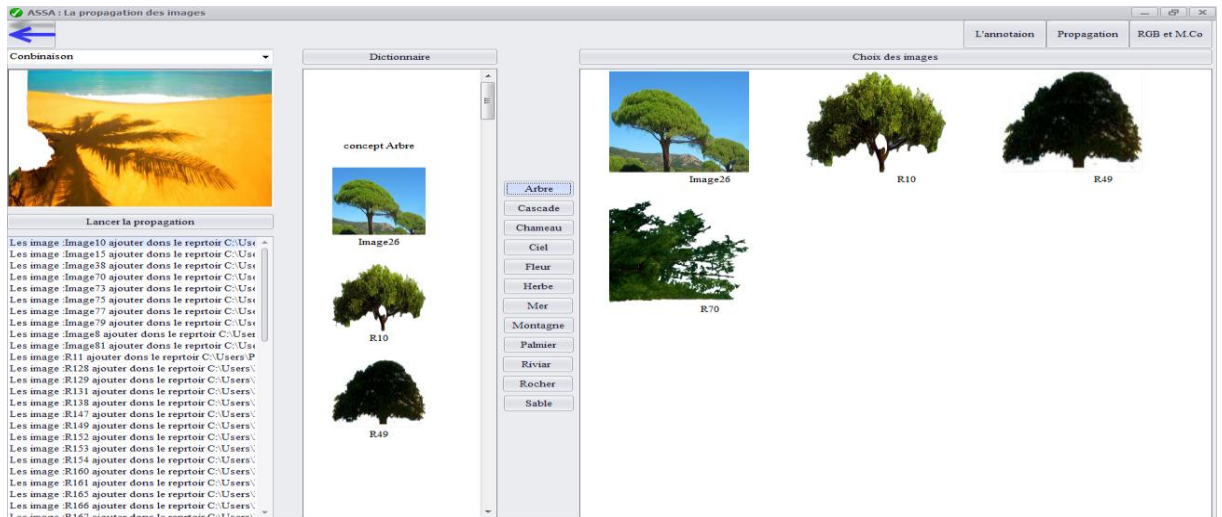


Figure - 47: Le contenu du concept arbre avant

## CHAPITRE V : Réalisation



**Figure - 48: Résultat Après mesure de similarité visuelle par la combinaison**

Après plusieurs expérimentations, nous avons tiré des constatations importantes :

- Pour certains concepts la mesure de similarité d'une seule caractéristique était suffisante. Et même quand nous avons appliqué la combinaison, le résultat n'a pas changé.
- Dans les cas de combinaison, on doit affecter des poids pour chaque caractéristique. Les poids se varient selon le concept, et la décision de ces valeurs est purement expérimentale.

Le tableau suivant résume nos résultats Après un nombre important d'expérimentations :

Les concepts	Caractéristiques utilisées	Le poids RGB	Le poids de la texture
Arbre	RGB	/	/
Cascade	Combinaison	0.7	0.3
Chameau	RGB	/	/
Ciel	Combinaison	0.8	0.2
Fleur	RGB	/	/
Herbe	Combinaison	0.6	0.4

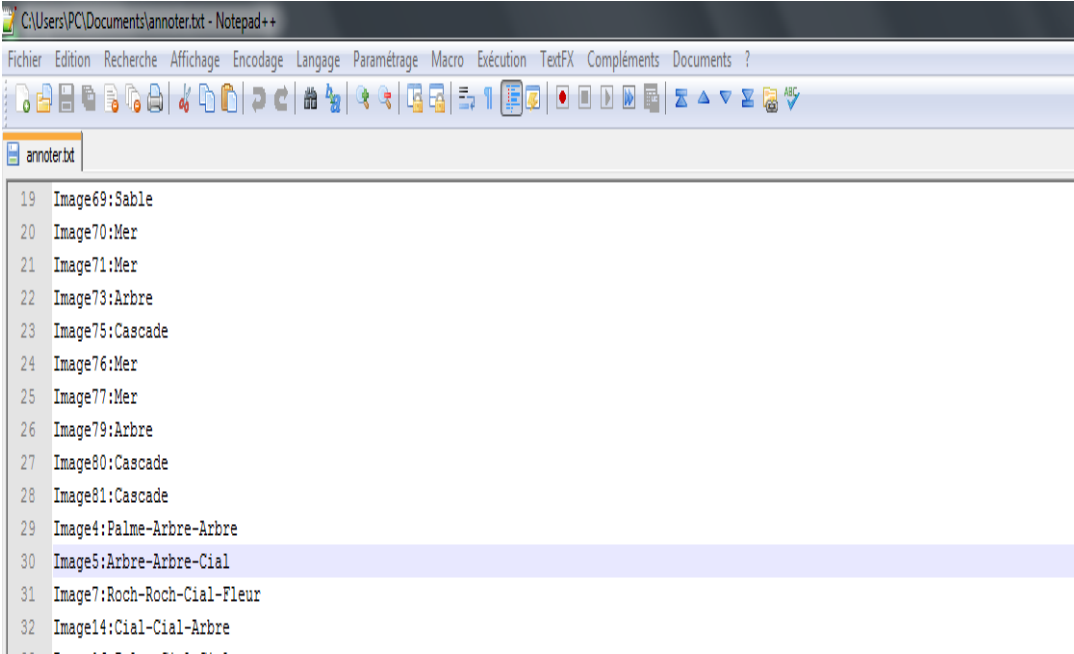
## CHAPITRE V : Réalisation

Mer	Combinaison	0.6	0.4
Montagne	RGB	/	/
Palmier	Combinaison	0.7	0.3
River	RGB	/	/
Roch	RGB	/	/
Sable	RGB	/	/

**Tableau2 : Sélection des caractéristiques et des poids pour chaque concept**

### 3.2.5.4. Association des concepts aux images

Il s'agit de récupérer en fonction des concepts des régions, les concepts des images. C'est une opération de correspondance simple.

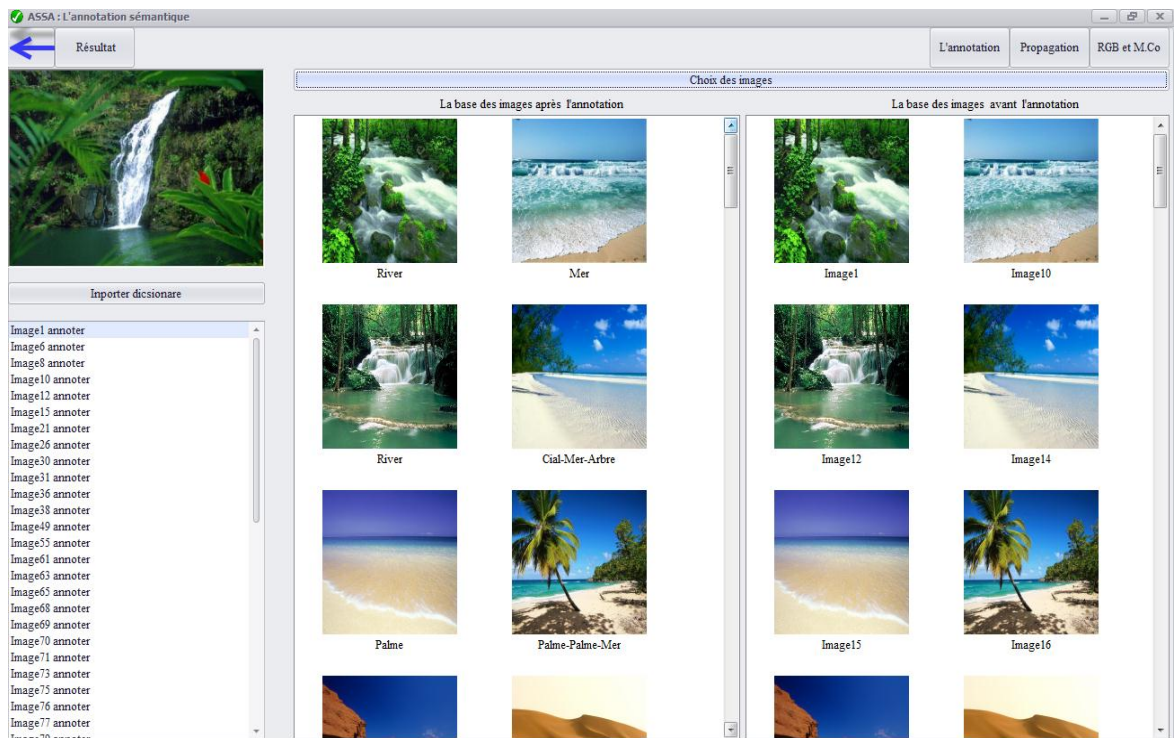


```
C:\Users\PC\Documents\annoter.txt - Notepad++
Fichier Edition Recherche Affichage Encodage Langage Paramétrage Macro Exécution TextFX Compléments Documents ?
annoter.txt
19 Image69:Sable
20 Image70:Mer
21 Image71:Mer
22 Image73:Arbre
23 Image75:Cascade
24 Image76:Mer
25 Image77:Mer
26 Image79:Arbre
27 Image80:Cascade
28 Image81:Cascade
29 Image4:Palme-Arbre-Arbre
30 Image5:Arbre-Arbre-Cial
31 Image7:Roch-Roch-Cial-Fleur
32 Image14:Cial-Cial-Arbre
```

**Figure - 49: Le fichier d'annotation résultat**

## CHAPITRE V : Réalisation

Une visualisation des annotations des images est faite :



**Figure - 50 : la base d'images avant et après l'annotation semi-automatique**

A ce stade nous arrivons à terminer notre système. Nous avons commencé par une base d'images non annotées, et nous arrivons à l'annoter avec des concepts feuilles de notre ontologie.

### 4. Conclusion

Nous avons présenté à travers ce chapitre l'implémentation de notre ontologie « Onto\_Nature » en utilisant l'éditeur protégé V3.3.4, ainsi que la réalisation de notre système ASSA qui s'articule principalement autour des étapes de filtrage, de segmentation, d'annotation manuelle d'un échantillon des régions, d'annotation automatique des régions restantes par mesure de similarité et propagation de concepts et enfin l'association des concepts aux images.

### Conclusion générale

Les bases d'images numériques connaissent un essor considérable. Leur facilité d'acquisition et de stockage les rendent très attractives pour des applications diverses. Il en résulte une production permanente et croissante d'images numériques dans différents domaines.

Par conséquent, et pour assurer une exploitation efficace de ces bases d'images par des machines, la description de telles bases est une tâche qui s'impose. Sinon, ces images peuvent demeurer pratiquement inexploitable et impossible de les retrouver et utiliser efficacement.

L'annotation des images est un domaine de recherche très actif. Elle consiste à désambiguïser et interpréter les images d'une façon qui assure leur exploitation adéquate par la machine pour des buts variés tels que : la recherche et la classification.

Dans le cadre de notre travail nous sommes intéressés par la proposition et l'implémentation d'une méthode d'annotation sémantique semi-automatique à base d'ontologie pour profiter de la richesse sémantique qu'elle présente.

Notre travail s'articule autour des phases principales suivantes :

1. D'abord, nous avons étudié les différentes méthodes d'annotations, ainsi que les différentes formes que peut avoir une annotation. Sur la base de cette étude, nous avons choisi l'annotation semi-automatique à base d'ontologie.

2. Par conséquent, nous avons construit une ontologie de domaine *Onto-Nature*. C'est une ontologie orientée terminologie sur le domaine de la nature. Elle est exploitée dans la phase d'annotation.

3. Nous avons implémenté une annotation sémantique semi-automatique, passant par les étapes suivantes :

- Segmentation manuelle de chaque image de la base en des régions
- Faire une annotation manuelle d'un échantillon de régions. De telle sorte que nous avons couverts tous les concepts sémantiques avec des apparences visuelles variées.
- Lancer l'annotation automatique pour le reste des régions par propagation de concepts sémantiques en se basant sur des mesures de similarité visuelles entre régions.
- Récupérer les concepts de chaque image.

## Conclusion Générale

---

Nous trouvons que notre choix d'une démarche d'annotation semi-automatique face aux méthodes manuelles et automatique est évident, car une méthode manuelle présente rapidement ses limites avec l'augmentation de la base d'images. Ainsi pour la méthode automatique il n'existe pas jusqu'à ce jour des travaux qui ont fait l'annotation d'une façon automatique proprement dite, car il faut toujours l'intervention de l'être humain au moins pour valider l'annotation.

De plus, l'utilisation d'une ontologie pour faire des annotations descriptives de contenu sémantique, est une solution très efficace vue à la richesse sémantique dans l'ontologie. Ces annotations sont développées actuellement autour du web sémantique, et semble l'approche la plus prometteuse pour partager, chercher et exploiter efficacement les ressources d'information.

Cependant, L'annotation sémantique, est difficile à automatiser et qui nécessite de nombreux essais pour évaluer la performance des algorithmes d'annotation. Cela due à la subjectivité, c'est-à-dire que chacun peut interpréter une ressource selon son point de vision. Pour cela, le développement des mécanismes d'annotation avec plusieurs niveaux d'abstraction est un axe important.

Pour cela, nous tirons plusieurs perspectives à notre travail :

5. Enrichir notre ontologie par d'autres concepts et relations sémantiques pour couvrir une variété des images de la nature avec une sémantique plus riche.
6. Après qu'on à arriver à développer un mécanisme d'annotation avec des concepts feuilles, il sera bénéfique de l'enrichir de tel sorte d'avoir plusieurs niveaux d'abstractions.
7. Faire une étude plus approfondie des méthodes de segmentation, et procéder la segmentation d'une façon automatique.
8. L'essai des autres caractéristiques visuelles pour la mesure de similarité.
9. Exploiter notre système pour servir d'autres travaux tel que la recherche d'images et la classification.



## BIBLIOGRAPHIE

- [Alain Boucher, 2005] **Alain Boucher**, indexation et recherche d'images par le contenu.2005
- [Amourache. F, 08] **Fouzia Amourache**, Construction d'une ontologie pour l'annotation des CVS/Offres d'emploi, 01/12/2008.
- [Baillie. J-C, 03] **Jean-Christophe Baillie**, "Traitement d'image et vision artificielle", ENSTA, 2003.
- [Bay. R, 12] **Bay Rokaia**, Une segmentation hybride d'images IRM cérébrales par combinaison de l'algorithme K-Means et Quad-Tree Génétique, 2012.
- [Ben Cheikh. N et al.2011] **Ben Cheikh Noura Ben Bezziane Rima**, la recherche d'images par la sémantique. Soutenu le /06 /2011.
- [Benoit. S, 08] **Benoit Scherrer**. Segmentation des tissus et structures sur les IRM cérébrales : agents markoviens locaux coopératifs et formulation bayésienne. Docteur en Mathématiques Appliquées et Informatique de l'INPG. Octobre 2008
- [Bon .J, 07] **Jacques BON**. " Image numérique". Document sous licence .2007.
- [Boucetta. Z, 08] **Boucetta Zouhel**, appariement sémantique des cvs/offres d'emploi dans le cadre du e-recrutement, 2008
- [Chargui. N, 2008] **Chargui Nabila**, une approche de mapping pour l'intégration des ontologies, 01/12/2008.
- [Do.M.C, 07] **DO Minh Chau**, Évaluation de la segmentation d'images , Institut de la francophonie pour l'informatique (IFI). 2007
- [Frédéric. F, 2002] **Frederic Furst**, l'ingénierie ontologique, No 02-07 Octobre 2002.
- [Ghandour. S, 10] **Sarah GHANDOUR**. "Segmentation d'images couleurs par morphologie mathématique: application aux images microscopiques". Thèse doctorat. Université de Toulouse. 2010.
- [Guediri. K, 12] **GUEDIRI Khaoula**, Une méthode floue adaptative (AF<sup>2</sup>PSO) pour la segmentation d'images IRM cérébrales, soutenu le juillet 2012.
- [Hacid. H, 06] **Hakim Hacid**, Annotation Semi-automatique de Grandes BD Images : Approche par Graphes de Voisinage, 2006.

**[Iggane.M et al, 09]** Mbark IGGANE, Driss MAMMASS, Ali KHENCHAF et Jean-Christophe CEXUS, Segmentation d'images sonar par matrice de co-occurrence , 2009

**[Jérémy.L et al, 08]** L.Jérémy et B.Christian. "*Segmentation d'images cérébrales*" .Rapport de recherche, Unité de recherche INRIA Rennes IRISA, Campus universitaire de Beaulieu, 18 Feb2008.

**[Kassous.S et al, 02]** S.KASSOUS, M. ALOIANI. "Segmentation d'images IRM cérébrale par coopération contour / région" .thèse d'ingénieur, institut d'électronique université de SAAD DAHLEB Blida, 2002.

**[Khouri. S, 2009]** Khouri Selma, Modélisation conceptuelle à base ontologique d'un entrepôt de données, 2009.

**[Mehanna. M et al, 2007]** Mahiddine Mehanna, Missoum Mehanna, conception et réalisation d'une ontologie dans le domaine des hydrocarbures pour la recherche d'information, institut National de formation en informatique (INI) Oued-Samar Alger, 2007.

**[Pham. S, 2004]** Phame Le Son, ontologie d'évènements vidéo pour un système automatique d'interprétation vidéo, 2004.

**[Roux. B, 2004]** Bernard le roux : recherche dans les bases de données images : une approche de description et une mesure de similarité combinant les aspects physiques et sémantiques de l'image. Soutenu le 22 novembre 2004

## WEBOGRAHIE

**[Site 1]** : <httpwww.map.toulouse.archi.frworkspanofformationimagenumimagenum.htm>, visité en : mai 2013

**[Site 2]** : [httpwww.memoireonline.com12093040m\\_La-liaison-automatique-des-plusieurs-images-percues-sur-un-scanner4.html](httpwww.memoireonline.com12093040m_La-liaison-automatique-des-plusieurs-images-percues-sur-un-scanner4.html), visité en : Mai 2013