

UNIVERSITE SAAD DAHLAB DE BLIDA1

Faculté des Sciences
Département Informatique

MEMOIRE DE MAGISTER

Spécialité : Informatique Répartie et Mobile (IRM)

PARALLELISATION DU PROCESSUS D'ANNOTATION
POUR LA RECHERCHE D'IMAGES DE PATIENTS
ATTEINTS DE RETINOPATHIE DIABETIQUE

Par

BOUKADOUM Omar

Présenté devant le jury composé de :

Mr A. GUESSOUM	Professeur, U. de Blida	Président
M ^r A. MEZIANE	Maître de recherche A, CERIST	Examineur
M ^{lle} N. BOUSTIA	Maître de Conférences B, U. de Blida	Examinatrice
M ^{lle} N. BENBLIDIA	Professeur, U. de Blida	Promotrice
M ^{me} S. OUKID-KHOUAS	Maître de Conférence A, U. de Blida	Co-promotrice

Blida, Mars 2016

RESUME

Dans notre travail, nous proposons une approche d'annotation automatique d'image par le contenu basée sur la structure arborescente. Notre approche est appliquée aux images rétinienne pour l'aide au diagnostic médical relatif à la rétinopathie diabétique (RD). Pour cela nous utilisons l'approche de l'arbre quaternaire (AQ) pour indexer une image. Chaque image de la base est représentée par un arbre quaternaire équilibré à trois niveaux dont les nœuds stockent des descripteurs visuels de la norme MPEG7. Pour l'annotation automatique d'image de la base, nous proposons de sauvegarder les mots clés associés aux pathologies (dans notre cas nous traitons les micro-anévrismes) dans le nœud racine de l'arbre. Dans l'objectif d'annoter une image requête, une distance de similarité sera calculée entre l'arbre quaternaire qui représente l'image requête et celles des arbres de la base.

Dans cette approche nous avons travaillé sur la base d'images de MESSIDOR [1] ainsi nous avons donné à l'expert la possibilité de choisir le descripteur adéquat (contour, texture, couleur), pour chaque niveau, par un simple clic sur ce dernier ainsi le choix du seuil. Notre système permet également d'afficher les résultats via une interface graphique efficace. Les résultats expérimentaux ont permis de montrer la robustesse offerte par notre système.

L'application de cette approche nécessite un temps élevé pour obtenir des résultats, de ce fait la technologie de parallélisation devient une nécessité pour gagner du temps.

La comparaison entre descripteurs est calculée parallèlement entre chaque deux nœud homologue. Ce calcul est indépendant les uns aux autres, tout en utilisant un réseau informatique qui contient un ensemble d'ordinateurs qui répartissent la charge de travail sur ces postes. Les résultats expérimentaux ont montré la réduction du temps par rapport à la méthode séquentielle.

Mots clés : annotation automatique, arbre quaternaire, descripteurs visuels, MPEG7, rétinopathie diabétique, parallélisme, réseau.

ABSTRACT

In our work, we propose an automatic image annotation approach by the content based on the tree structure. Our approach is applied to the retinal images to aid in medical diagnosis on diabetic retinopathy (DR). For this we use the approach of the quaternary tree (AQ) for indexing an image. Each image of the base is represented by a balanced quaternary tree with three levels whose nodes store visual features of the MPEG7 standard. For automatic annotation of the base image, we propose to save the keywords associated pathologies (in our case we treat micro-aneurysms) in the root node of the tree. With the aim to annotate a query image, a similarity distance is calculated between the quad-tree representing the query image and those of the base of trees.

In this approach we worked on the basis of images of MESSIDOR [1], and we gave the expert the opportunity to choose the appropriate descriptor (outline, texture, color) for each level, with a single click on this and the last choice of threshold. Our system also displays the results via an efficient graphical interface. The experimental results have shown the robustness offered by our system.

The disadvantage of the application of this approach is time high implementation, to achieve results in a reasonable time in real-time, we believe that parallelization technology is indispensable in this case to save time.

The comparison of descriptors will be computed in parallel course between each two peer nodes because all these calculations are independent of each other using a computer network that contains a set of computers to distribute the workload on those positions.

Keywords: automatic annotation, quaternary tree, visual features, MPEG7, diabetic retinopathy, parallelism, network.

ملخص

في علنا هذا، اقترحنا نهجا للشرح الألي لصور مبني على أساس المحتوى ونستند على الهيكل الشجري. تقنيتنا تطبق على الصور التي تحتوي على الشبكيات لتشخيص المرض الخاص بالشبكية السكري، لهذا نستعمل الشجرة الرباعية لفهرسة الصورة، كل صورة من قاعدة البيانات تمثل بشجرة رباعية متوازنة ذات 3 مستويات بحيث كل عنصر من الشجرة يقوم بتخزين الواصفات المعتدة من طرف MPEG-7.

للشرح الألي للصورة، نقترح تخزين الكلمات الدالة على المرض (حالة علاج مرض microanévrisme) في عنصر المستوى الأول للشجرة. يتم حساب المسافة للمقارنة بين صورتين من قبل اثنين من الأشجار الرباعية.

استعملنا قاعدة معطيات MESSIDOR كمرجع وأتحننا للمختص امكانية اختيار الواصف المناسب (اللون، النسيج، الشكل) لكل مستوى وكلك ادخال قيمة العتبة، ولقد اظهرت النتائج التجريبية مدى متانة، قوة ونجاح هذه التقنية التي تقدم ذكرها. من بين سلبيات هذه التقنية هو الوقت الكبير الذي يستغرقه البرنامج لاعطاءنا النتائج في الوقت الحقيقي وهذا كان علينا ادخال نظام التوازي كوسيلة حتمية لربح الوقت، بحيث أننا قمنا بحساب الواصفات بين كل عنصرين متماثلين من الشجرتين التي نريد مقارنتهما وهذا بالتوازي باستعمال شبكة الاعلام الألي التي تحتوي على مجموعة من الحواسيب لتخفيف الضغط وتقسيم العمل فيما بينها.

كلمات مفتاحية: شجرة رباعية، MPEG-7، الواصفات البصرية، التوازي، شبكة، الشبكية السكري

REMERCIEMENTS

Avant tout, nous remercions DIEU le tout-puissant de nous avoir donné le courage, la volonté, la patience et la santé durant toutes ces années d'études. C'est grâce à lui que ce travail a pu être réalisé.

Je tiens à exprimer mes plus vifs remerciements à Mlle BENBLIDIA Nadjia ainsi Mme Oukid-Khouas Saliha pour leur patience, leurs encadrement, leurs disponibilité, leurs apport à ce mémoire, leurs corrections et tous les conseils judicieux qu'elles n'ont cessé de me prodiguer.

Le remercie aussi tous les enseignant qui, dans le cadre de magister, ont mis à ma disposition leurs connaissances chacun dans sa spécialité.

Je tiens à remercier mon ami KHALFI Ali qui m'a aidé dans ce mémoire.

Je tiens à remercier tous les membres du jury d'avoir acceptés d'examiner ce mémoire.

Je remercie mes parents, mes frères, sœurs et toute ma famille pour leur soutien dans ce projet et tous ceux qui ont pu nous aider dans la réalisation de ce travail sans oublier les amis et les collègues pour leur présence à nos côtés tout au long de cette étude.

LISTE DES FIGURES

Figure 1.1	Chaine de vision par ordinateur.	16
Figure 1.2	Une coupe du globe oculaire.	21
Figure 1.3	Les différentes zones de la rétine.	24
Figure. 1.4	Présence des micro-anévrismes et hémorragies dans les images couleur de la rétine ; (a) : micro-anévrismes montrés par une flèche ; (b) : hémorragies.	27
Figure 1.5	Présence des AMIR et une anomalie veineuse rétinienne.	27
Figure 1.6	RD non proliférante modérée : hémorragie en tache peu nombreuses.	29
Figure 1.7	RD non proliférante sévère : nombreuses hémorragies en taches sur toute la périphérie.	29
Figure 1.8	Néovaisseau pré-rétinien.	29
Figure 1.9	Néovaisseaux pré-papillaires.	29
Figure 1.10	Œdème maculaire focal localisé, entouré d'exsudats.	30
Figure 2.1	Lac du Passet	37
Figure 2.2	Un exemple de recherche d'images dans Google.	38
Figure 2.3	Deux découpages différents de l'image. (a) Découpage en grille de l'image, (b) Découpage par la segmentation.	41
Figure 2.4	Une requête par esquisse dans QBIC.	43
Figure 2.5	Une requête par image dans Blobworld.	44
Figure 2.6	Requête par image exemple dans le système CBIR.	45
Figure 2.7	Description de radiographie cérébrale par le contenu sémantique.	48
Figure 2.8	Un exemple de retour de pertinence.	51
Figure 2.9	Interface pour annoter une image, Un arbre de mots clés (à gauche), image à annoter (à droite).	52
Figure 2.10	Quelques images annotées avec ses mots clés choisis.	53

Figure 3.1	Architecture générale d'une approche d'annotation automatique d'image.	61
Figure 3.2	Etapes d'extraction du contenu Visuel d'image.	62
Figure 3.3	Extraction du contenu visuel d'une image.	63
Figure 3.4	Exemple de segmentation d'image.	64
Figure 3.5	Les approches de segmentation d'image.	64
Figure. 4.1	Représentation d'une image par l'arbre quaternaire.	71
Figure 4.2	Structure d'arbre quaternaire équilibré proposée pour notre système d'annotation automatique.	75
Figure 4.3	Annotation automatique d'image par arbre quaternaire.	76
Figure 4.4	Machine multiprocesseurs à mémoire partagée. p représente un processeur indépendant.	83
Figure 4.5	Modélisation d'une machine parallèle MIMD à mémoire distribuée. Chaque nœud est composé d'un processeur et d'une mémoire locale. Les nœuds communiquent entre eux en envoyant et en recevant des messages via le réseau de communication.	84
Figure 4.6	Architecture du système SGI Origin. Deux processeurs sont connectés via un « hub » à une mémoire partagée pour former un « nœud ». Plusieurs nœuds sont ensuite connectés en utilisant un réseau d'interconnexion pour former le système global.	85
Figure 4.7	Grappe de PCs du projet iCluster (HP, INRIA, ID-IMAG) composée de 226 mono-processeurs Pentium III à 733 MHz avec 256 Mo de mémoire, inter-connectés via un réseau de 100 Mbit/s . (A droite) Grappe du projet iCluster2 composée de 104 bi-processeurs Itanium-2, 64 bits à 900 MHz avec 3 Go de mémoire, inter-connectés par un réseau Myrinet.	86
Figure 4.8	Schéma du réseau informatique.	87
Figure 4.9	Modélisation de notre proposition de parallélisme.	87
Figure 4.10	Modèle d'exécution Maître/Esclave	89
Figure 5.1	Exemple d'images rétiniennes de la base de données MESSIDOR (a) : image saine (normale) ; (b) : image pathologique.	94
Figure. 5.2	Résultats de recherche pour l'image requête pathologique (avec microanévrisme).	96

LISTE DES TABLEAUX

Tableau 1.1	Lésions de la Rétinopathie Diabétique.	26
Tableau 1.2	Classification de la Rétinopathie Diabétique.	30
Tableau 5.1	Résultats de l'évaluation des images.	95
Tableau 5.2	Temps d'exécution général séquentiel entre deux arbres.	97
Tableau 5.3	Temps d'exécution dans le cas parallèle.	98
Tableau 5.4	Evolution du nombre de machines par rapport au niveau de détail.	99

TABLE DES MATIERES

RESUME	
ABSTRACT	
ملخص	
DEDICACES	
REMERCIEMENTS	
LISTE DES FIGURES	
LISTE DES TABLEAUX	
TABLE DES MATIERES	
INTRODUCTION GENERALE	11
1. CONTEXTE BIOMEDICAL ET VISION ASSISTEE PAR ORDINATEUR	13
1.1 Introduction	13
1.2 Vision Humaine/Vision Assistée par Ordinateur	13
1.3 Principaux Traitement	16
1.4 Contexte Biomédical	19
1.5 Conclusion	32
2. RECHERCHE D'IMAGES	33
2.1 Introduction	33
2.2 Domaines d'application de la recherche d'images	33
2.3 Bases d'images	34
2.4 Approches de représentation et de recherche d'images	35
2.5. Conclusion	54
3. ANNOTATION AUTOMATIQUE D'IMAGES	55
3.1 Introduction	55
3.2 Définition de l'annotation d'images	55
3.3 Etat de l'art sur l'annotation d'image	57
3.4 Architecture générale d'une approche d'annotation automatique d'image	59
3.5. Modèles d'annotation automatique d'images à partir du contenu visuel	65
3.6. Conclusion	68
4. APPROCHE PARALLELE D'ANNOTATION AUTOMATIQUE D'IMAGES	69
4.1 Introduction	69
4.2. Concept de base de l'arbre quaternaire	69
4.3. Similarité d'images représentées par des arbres quaternaires	71
4.4. Approche proposée en annotation automatique d'image basée sur l'arbre	73
4.5 Parallélisme d'annotation automatique d'image proposé	80

	10
4.6 Conclusion	90
5. IMPLEMENTATION ET TESTS	91
5.1 Introduction	91
5.2 Description de la base MESSIDOR	91
5.3 Implémentation de l'arbre quaternaire	91
5.4 Experimentations et resultats	93
5.5. Discussion	97
5.6. Conclusion	99
CONCLUSION GENERALE ET PESPECTIVES	100
REFERENCES BIBLIOGRAPHIQUE	102

INTRODUCTION GENERALE

Les images ont toujours été utilisées en médecine pour l'enseignement et le diagnostic. De nos jours, les systèmes d'imagerie médicale produisent de plus en plus d'images numériques dans tous les domaines de la médecine : le visible, l'ultrason, la tomographie à rayons X, l'IRM, l'imagerie nucléaire, etc... Ces images sont très intéressantes à des fins de diagnostic : elles sont directement liées à la pathologie du patient ainsi qu'à son histoire médicale.

Nous nous intéressons plus particulièrement dans notre travail à l'annotation automatique des images rétinienne pour la détection des pathologies de la rétinopathie diabétique (RD), une maladie de œil très savère et répandue.

La rétinopathie diabétique (RD) est une maladie à forte prévalence chez les diabétiques. Il existe des traitements pour prévoir sa progression et réduire le risque de cécité. L'efficacité du traitement est plus grande lorsque le diagnostic est posé tôt et qu'il est administré en temps opportun. Après analyse des données probantes, des expériences existant ailleurs et du contexte actuel en Algérie, la mise en place d'un programme de dépistage de la rétinopathie diabétique est pertinente et faisable. Un tel programme vise à déceler les personnes qui présentent une rétinopathie à haut risque, car elles peuvent bénéficier d'un traitement préventif. Avant d'établir des scénarios réalistes et d'estimer leurs coûts, il faudra définir, entre autres, l'organisation du dépistage dans le cadre de la prise en charge globale actuelle des patients diabétiques, le partage des activités de dépistage entre les professionnels, et la formation de professionnels qui devront assurer la qualité des photos. [1]

Depuis une dizaine d'années, de nombreuses études ont été menées au niveau mondial pour développer des techniques automatiques de dépistage et de suivi de la RD.

Ces systèmes, basés sur le traitement automatisé des images numériques, comprennent principalement des outils de détection des lésions élémentaires qui caractérisent la RD comme par exemple : (micro-anévrisme, exsudats, hémorragies,..), et des outils d'indexation et d'annotation automatiques dans des bases d'images déjà annotées, dans ce mémoire nous nous intéressons aux outils de la deuxième catégorie.

Ce mémoire est abordé selon l'organisation suivante :

Chapitre 1 : Contexte biomédical

Ce chapitre préliminaire décrit la vision Humaine/Vision Assistée par Ordinateur, Les principaux Traitement d'Images et le cadre biomédical du projet : l'anatomie de l'oeil particulièrement la rétine, les différentes techniques d'exploration de la rétine notamment l'angiographie rétinienne à la fluorescéine, la description de la pathologie oculaire à laquelle nous nous intéressons, la rétinopathie diabétique.

Chapitre 2 : Recherche d'images

Nous rappelons dans ce chapitre à un état de l'art des différentes approches de représentation et de recherche d'image connus dans la littérature, ensuite les formes de requêtes dans les systèmes de recherche d'images dans chaque approche ainsi leurs limites.

Chapitre 3: Annotation Automatique d'Image

Dans ce chapitre on va faire un survol sur des modèles et des méthodes déjà utilisé dans l'annotation. Dans ce chapitre, nous rappelons tout d'abord l'annotation d'images et leurs types. Ensuite, nous allons introduire quelques travaux existants sur l'annotation automatique d'images puis nous présentons une architecture générale d'une approche d'annotation automatique d'image et leurs détails, ensuite nous présentons quelques modèles d'annotation automatique d'images qu'il existe dans la littérature.

Chapitre 4 : Approche parallèle d'annotation automatique d'image

Dans ce chapitre, nous décrivons l'approche adoptée pour atteindre les objectifs fixés Nous exploitant les descripteurs visuels de la norme MPEG7 (Moving Picture Experts Group version 7) de couleur, texture et de forme. Puis nous introduisons la notion de parallélisme dans l'approche d'annotation en utilisant le réseau informatique pour diviser la charge de travail (le calcul) sur les ordinateurs du réseau qui sont de nombre 21 machines.

Chapitre 5: Implémentation et Tests

Ce chapitre présente l'environnement sur lequel on a mené notre travail, les tests effectués sur une base d'images rétinienne annotées, nous avons discuté les différents résultats obtenus par la technique proposée. Nous terminons par une conclusion générale et fixe les perspectives envisageables.

CHAPITRE 1

CONTEXTE BIOMEDICAL ET VISION ASSISTEE PAR ORDINATEUR

1.1 Introduction

Ce chapitre présente les aspects médicaux, nécessaire pour comprendre l'importance de l'annotation des images du fond d'œil. Nous commençons tous d'abord, par des notions générales sur la vision assistée par ordinateur (VAO) ainsi des principaux traitements d'images ensuite en passe au domaine médical de notre travail c'est la Rétinopathie Diabétique.

1. 2 Vision Humaine/Vision Assistée par Ordinateur

L'avènement des robots de la troisième génération capable de percevoir leur environnement et de réagir à celui-ci a fait faire un bond en avant aux théories et techniques de la perception. Parmi celles-ci, la vision s'est affirmée comme un domaine de recherche privilégié, et ce pour plusieurs raisons. De tous nos sens, la vue est celui qui offre la plus grande flexibilité et nous apporte la plus grande quantité d'information sur ce qui nous entoure du fait de la multiplicité de ses dimensions : spatial, énergétique, temporelle.

La perception de notre espace, la reconnaissance des objets qui le compose et la détection du mouvement de ceux-ci sont des données essentielles pour l'humain et sont toutes transmises par le biais de ses yeux.

Le traitement numérique des images va mettre en œuvre deux types d'approches principaux :

- L'amélioration d'images pour la visualisation et éventuellement l'interprétation manuelle par un expert humain.

- La vision par ordinateur qui consiste à réaliser des opérations de perception, d'interprétation automatique par ordinateur, de façon analogue au système de perception visuelle humain.

1.2.1 Perception Humaine

L'œil doit d'abord obtenir une image du monde extérieur parfaitement au point sur la rétine, ce qui demande un système optique élaboré (composé de la cornée, du cristallin et du diaphragme irien). Ensuite le système nerveux (rétine, voies visuelles et cerveau) doit transmettre fidèlement cette image depuis l'œil jusqu'au cerveau. Enfin, ce dernier doit réélaborer les informations reçues pour les faire émerger sous la forme d'une image.

1.2.2 Vision assistée par ordinateur (VAO)

A ces considérations s'ajoute un facteur non négligeable qui est la possibilité de l'homme de créer un être à son image, depuis les premiers automates jusqu'aux systèmes de vision stéréoscopiques qui utilisent deux caméras juxtaposées de manière analogue aux deux yeux de l'être humain, ce qui permet de percevoir la distance de l'objectif à l'objet examiné.

L'arrivée sur le marché d'ordinateurs de plus en plus performants a rendu possible ce qui n'était qu'une vue de l'esprit il y a quelque dizaines d'années. Des progrès de la technique est née la vision par ordinateur (VAO) dans laquelle vient s'intégrer le traitement d'images.

La conception, des systèmes de vision pour l'interprétation automatique des scènes suscite un intérêt particulier et de plus en plus croissant. C'est dans ce cadre, que de nombreux chercheurs se sont penchés sur ce problème et ont développé plusieurs approches en vue de la conception d'un système de vision adéquat qui puisse décrire l'importance des quantités d'informations contenues dans l'image en recherchant des indices visuels ou des primitives pertinentes permettant de la représenter sous une forme plus condensée et facilement exploitable par la machine : c'est le domaine de l'analyse d'images.

La vision assistée par ordinateur (VAO) et, à travers elle, le traitement d'images ont connu ces dernières années une très forte expansion dans divers domaines tels que l'industrie, la robotique, l'espace, le médical, le militaire, etc. cela peut s'expliquer par la part prépondérante que représente la vision dans l'activité humaine et par le besoin d'améliorer cette l'exécution autonome de certaines tâches (voir la figure 1.1).

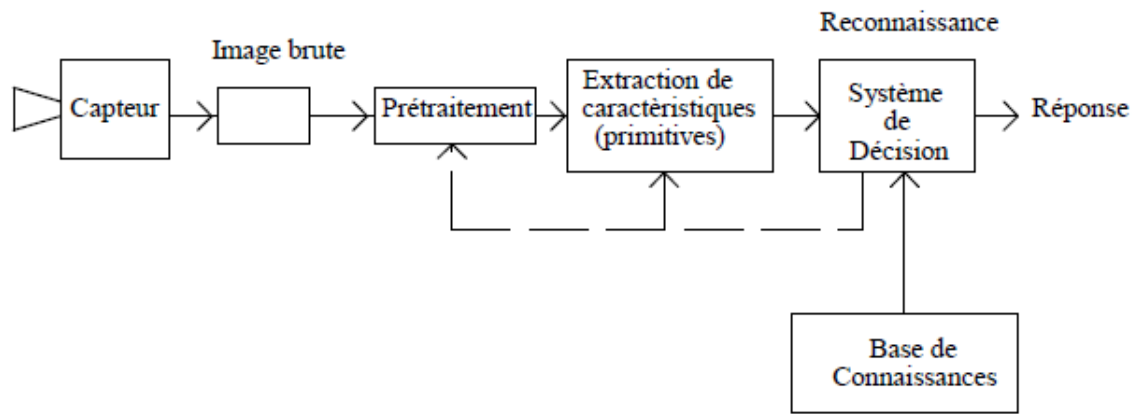


Figure 1.1 : chaîne de vision par ordinateur. [1]

Capteur : caméra vidéo, CCD,..., plaque photosensible aux rayonnements (Gama,X,...)

Prétraitement : amélioration de l'image (renforcement du contraste, réduction du bruit,...)

Extraction de caractéristiques : contours, régions, angles, segments de droites,...

Les caractéristiques ou primitives : sont compatibles avec la représentation qui en est donné dans la Base de connaissances (mémoire).

Système de Décision : utilise des méthodes de Reconnaissance Des Formes (RDF) communes à d'autres disciplines (reconnaissance de la parole,...)

Analogie : le cerveau corrige l'image floue et déformée issue de l'œil (même chose pour les oreilles !) et effectue de nombreux prétraitements (occupant ainsi des ressources du cerveau, et, chez des personnes âgées, cette consommation de ressources peut se faire au détriment de l'attention).

1.3 Principaux Traitement d'Images

Le traitement d'images est une discipline de l'informatique et des mathématiques appliquées qui étudie les images numériques et leurs transformations, dans le but d'améliorer leur qualité ou d'en extraire de l'information.

Il s'agit d'un sous-ensemble du traitement du signal dédié aux images et aux données dérivées comme la vidéo (par opposition aux parties du traitement du signal consacrées à d'autres types de données : son et autres signaux monodimensionnels notamment), tout en opérant dans le domaine numérique (par opposition aux techniques analogiques de traitement du signal, comme la photographie ou la télévision traditionnelles).

Dans le contexte de la vision artificielle, le traitement d'images se place après les étapes d'acquisition et de numérisation, assurant les transformations d'images et la partie de calcul permettant d'aller vers une interprétation des images traitées. Cette phase d'interprétation est d'ailleurs de plus en plus intégrée dans le traitement d'images, en faisant appel notamment à l'intelligence artificielle pour manipuler des connaissances, principalement sur les informations dont on dispose à propos de ce que représentent les images traitées (connaissance du « domaine »).

La compréhension du traitement d'images commence par la compréhension de ce qu'est une image. Le mode et les conditions d'acquisition et de numérisation des images traitées conditionnent largement les opérations qu'il faudra réaliser pour extraire de l'information. En effet, de nombreux paramètres entrent en compte, les principaux étant :

- la résolution d'acquisition et le mode de codage utilisé lors de la numérisation, qui déterminent le degré de précision des éventuelles mesures de dimensions,
- les réglages optiques utilisés, (dont la mise au point) qui déterminent par exemple la netteté de l'image,
- les conditions d'éclairage, qui déterminent une partie de la variabilité des images traitées,
- le bruit de la chaîne de transmission d'image.

1.3.1 Quelques exemples concrets de traitement d'images [1]

- Contrôle de présence/absence. Sur des chaînes de production, on vérifie en bout de chaîne avec une caméra vidéo la présence d'une pièce dans un ensemble plus complexe. Pour cela bien souvent il suffit de faire un simple seuillage dans une région spécifique.
- Contrôle du niveau de maturation des fruits sur une chaîne de conditionnement. Il s'agit de reconnaître à la couleur et à la texture du fruit son degré de maturité et donc la catégorie sous laquelle il sera emballé puis vendu.
- Construction et correction de cartes géographiques d'après des images satellites ou des images aériennes. On recale d'après des informations topographiques les images reçues, puis on les met sur la carte en correspondance avec les informations trouvées dans l'image : voies de communication, voies et plans d'eau, parcelles agricoles...
- Surveillance et évaluation de la production agricole. Il est possible de déterminer le degré de maturation des cultures, la quantité d'eau nécessaire pour l'irrigation, le rendement moyen... On peut ainsi établir des prévisions à large échelle de la récolte à venir.
- Reconnaissance de l'écriture. La reconnaissance de l'écriture manuscrite progresse de jour en jour. Elle est suffisamment opérationnelle pour que la majorité des adresses, même manuscrites, soient reconnues automatiquement sur le courrier postal.
- Recherche d'image par le contenu. L'objectif de cette technique est de rechercher, parmi une base de données d'images, les images similaires à une image exemple, ou ayant certaines caractéristiques, par exemple rechercher toutes les images comportant un vélo.
- Analyse de la vidéo. L'objectif de cette technique devenue une discipline depuis les années 2000 (lorsque la puissance des processeurs peu onéreux et en particulier des PC a permis des traitements puissants en temps réel) est d'interpréter les faits observés à l'image afin de signaler ou d'enregistrer des faits marquants. Le plus souvent, la caméra est fixe et observe les mouvements d'une scène. Les applications sont nombreuses : protection des biens (détection d'intrusion, détection d'objet abandonné

ou déposé...), identification (biométrie faciale), Sécurité des personnes (détection de chutes de personnes, franchissement de rambardes, ...), animations (planchers animés selon les mouvements des danseurs en boîte de nuit), détection de feux (industriel, forêts, tunnels, ...), surveillance de tunnels (comptage, mesure de vitesse, détection de fuites/anomalies dans les plafonds), surveillance de tuyaux et autres process industriels, ...

- Segmentation et suivi de cellules vivantes en microscopie. Cela permet d'analyser le comportement d'une population de cellules et ainsi de détecter certaines anomalies.
- Analyse et authentification de tableaux. L'étude des niveaux des couleurs des pigments et des vernis permet une analyse approfondie des œuvres. Il est ainsi possible de voir les restaurations successives et d'identifier les faux.

1.3.2 Principales étapes [2]

Il n'existe pas de méthode de traitement d'images générale à tous les domaines d'application possibles. Il faut en général employer des algorithmes spécifiques. Ces derniers sont souvent des combinaisons de techniques classiques (segmentation, classification, reconnaissance de frontières, etc.). De manière schématique, toute méthode de traitement d'images comprend 4 étapes majeures :

- Prétraitement des images,
- Amélioration des images,
- Analyse des images,
- Interprétation des images.

1.4 Contexte Biomédical

1.4.1 Introduction

Ce mémoire porte sur le processus d'annotation pour la recherche d'images de patients atteints de rétinopathie diabétique (d'images rétinienne). L'objectif est le développement d'un outil qui servira à une aide au diagnostic en ophtalmologie, notamment pour le dépistage de la Rétinopathie Diabétique, une maladie particulièrement sévère et répandue. Dans ce chapitre préliminaire nous décrivons le cadre biomédical de ce mémoire.

Nous commençons par une description de l'anatomie de l'œil en nous focalisant sur la rétine pour faciliter la compréhension de ce notre travail. Ensuite, Nous décrivons brièvement les techniques d'exploration du fond d'œil (rétine), et enfin nous aborderons les pathologies liées à la rétine et nous décrivons la Rétinopathie Diabétique, sujet de ce mémoire, en s'intéressant au cours de cette maladie, ces stades de complications et ces symptômes.

1.4.2 Anatomie de l'œil

L'œil est l'organe de la vision. Il est de faible volume (6.5 cm³), il pèse 7 grammes, il a la forme d'une sphère d'environ 24 mm de diamètre, complétée vers l'avant par une autre sphère de 8 mm de rayon, la cornée. La paroi du globe oculaire est formée de 3 tuniques : *la tunique fibreuse ; la tunique uvéale ; la tunique nerveuse*

Ces tuniques enferment des milieux transparents : l'humeur aqueuse ; le cristallin ; le corps vitré. Voici une coupe du globe oculaire afin de situer chacune des parties : (figure 1.2)

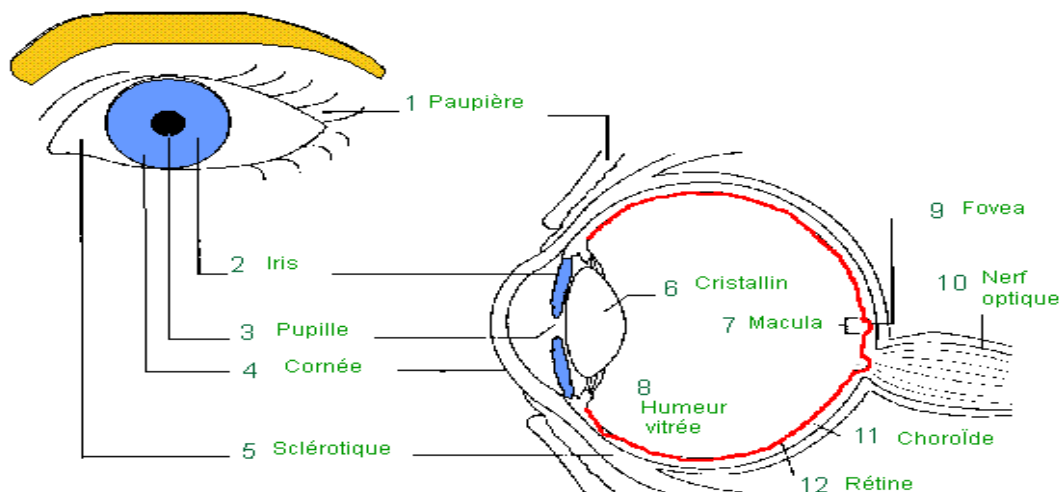


Figure 1.2 : une coupe du globe oculaire

La tunique fibreuse : La tunique fibreuse, externe, se compose de la sclérotique opaque en arrière et de la cornée transparente en avant.

- **la sclérotique** : C'est une membrane rigide qui donne sa forme à l'œil. Devant, elle devient la cornée. Membrane fibreuse, résistante, blanche chez l'adulte, bleue chez l'enfant, jaunâtre chez le vieillard et qui entoure et protège l'œil à l'extérieur tout en maintenant sa forme. Comme elle est inextensible, les variations de la pression intra-oculaire ne modifient pas le volume de l'œil. La sclérotique forme ce que l'on appelle couramment "le blanc de l'œil". Elle est traversée par un grand nombre de petits canaux (artères, nerfs, veines) et, à l'arrière, par une ouverture où passent les fibres du nerf optique et qui s'appelle la lame criblée.
- **la cornée** : La cornée constitue la lentille principale du système optique oculaire ; pour que ce tissu puisse remplir sa fonction il doit être transparent et tout concourt à assurer cette transparence. La cornée est enchâssée dans l'ouverture antérieure de la sclérotique ; elle est la structure qui a la plus grande sensibilité tactile du corps humain. Elle se compose de 5 couches, son épaisseur totale est d'environ 0,5mm au centre et 0,8mm à la périphérie. Le rayon de courbure de la face antérieure (face en contact avec l'air) est égal à 7,8mm en moyenne. La frontière entre cornée et sclérotique, appelée limbe, est une zone semi-transparente qui a la particularité

d'adhérer à la conjonctive, fine membrane qui couvre la face interne des paupières et la portion antérieure de la sclérotique. Tandis que la cornée est avasculaire, le limbe est richement innervé et vascularisé.

La tunique uveale : La tunique uvéale, dite aussi uvée, se compose de 3 éléments : l'iris en avant, le corps ciliaire, la choroïde en arrière.

- ***L'iris*** : L'iris est la portion la plus antérieure de l'uvée; il est de structure pigmentée, donnant sa couleur à l'œil et percé d'un trou, la pupille. La contraction ou la dilatation réflexe règlent la quantité de lumière pénétrant dans l'œil par la pupille. Il est situé dans l'humeur aqueuse entre la cornée et le cristallin, séparant ainsi la chambre antérieure de la chambre postérieure de l'œil. L'iris est innervé par des fibres du système nerveux autonome qui activent les muscles dilatateurs et sphincters responsables de la dilatation et de la constriction pupillaires.
- ***le corps ciliaire*** : Le corps ciliaire est représenté par un épaississement de l'uvée situé en couronne derrière l'iris. C'est une série d'environ 80 fins processus radiaires contenant des fibres musculaires lisses et des vaisseaux sanguins.
- ***la choroïde*** : La choroïde est une couche vasculaire, située entre l'épithélium pigmentaire et la sclérotique. C'est le tissu nourricier de l'œil. Il apporte l'oxygène et les nutriments dont les cellules ont besoin pour vivre.

La tunique nerveuse : La tunique nerveuse se compose de la rétine.

- ***la rétine*** : La rétine est un tissu transparent qui est responsable de la transformation de la lumière en impulsions électriques. Il s'agit d'un tissu neuronal très fin et hypersensible qui fait partie du système nerveux central : elle peut être vue comme un prolongement direct du cerveau dans l'œil. Elle recouvre la choroïde du côté intérieur.

La rétine est composée de deux feuillets, le feuillet extérieur (épithélium pigmentaire), adjacent à la choroïde, d'environ 10 μm d'épaisseur et le feuillet intérieur (rétine neuro-sensorielle) d'une épaisseur d'environ 200 μm qui, de son côté, peut être divisé en plusieurs couches. L'épithélium pigmentaire sert de chambre noire et à des fonctions majeures dans le métabolisme de la rétine neuro-sensorielle. Le feuillet neuro-sensoriel contient les photorécepteurs qui sont responsables de la transformation de la lumière en signal neuronal.

Ce signal est ensuite transmis au cerveau ou il est décodé pour permettre enfin la formation de l'image. Il existe deux types différents de récepteurs : les bâtonnets pour la vision en noir et blanc (la vision nocturne) et les cônes pour la vision en couleur (vision en éclairage diurne). La sensibilité des bâtonnets est plus élevée que celle des cônes et elle dépend fortement de la longueur d'onde de la lumière incidente (elle est maximale pour une longueur d'onde d'environ 500 nm). Parmi les cônes, il existe trois types différents, qui sont chacun responsable de la réception de la lumière dans une région spectrale différente : il y a des récepteurs pour les ondes courtes, moyennes et longues. La couleur est le résultat de la superposition des trois impulsions produites par les trois récepteurs [1]. On a beaucoup plus de bâtonnets (130 millions) que de cônes (6-7 millions). Le diamètre des cônes est beaucoup plus petit que celui des bâtonnets. Plus on s'éloigne de la partie centrale, plus les cônes se font rares et leur diamètre augmente [2].

La macula et la fovéa

Dans la zone elliptique centrale se trouve le maximum de cônes. Cette zone qui permet une vision très précise, mesure 3 mm dans le grand axe et 2 mm dans le petit axe. Elle s'appelle la macula. La macula, tâche jaune, apparaît située au centre du pôle postérieur comme une fine excavation [2].

La fovéa est une région de la rétine située dans la macula, près de l'axe optique de l'oeil. Cette région est de la plus haute importance pour la vision. C'est elle qui donne la vision la plus précise, en éclairage diurne, c'est-à-dire pendant la journée. Quand nous fixons un objet, nous tournons les yeux de façon à aligner l'image sur cette partie de la rétine. La fovéa est la partie centrale de la macula. Elle mesure 1300 à 1500 microns. Elle contient 400 000 cônes. Les cônes ont besoin de plus de lumière que les bâtonnets pour être excités. Les cônes réagiront plus en éclairage diurne qu'en éclairage nocturne. Les bâtonnets ont besoin de beaucoup moins de lumière pour réagir, ils assurent la vision nocturne [2].

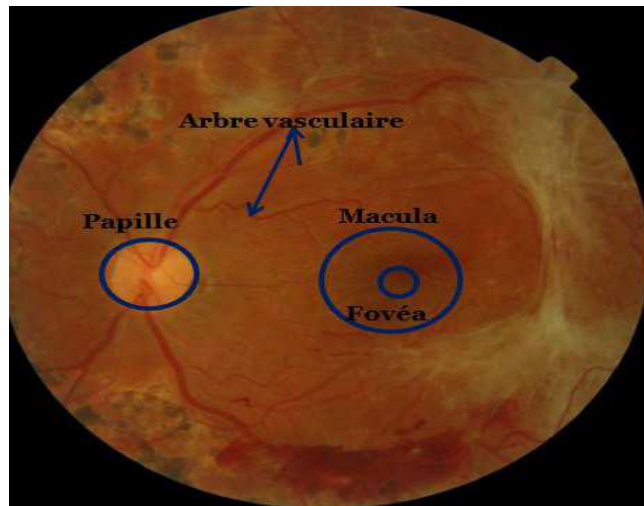


Figure 1.3 : Les différentes zones de la rétine

Le paragraphe suivant sera dédié à la description de la rétinopathie diabétique. Une maladie très sévère de la rétine, c'est une affection ophtalmologique considérée comme la complication la plus redoutable du diabète.

1.4.3 Epidémiologie

De grandes études épidémiologiques ont permis de mieux cerner l'incidence des complications rétiniennes de la rétinopathie diabétique (RD).

On estime qu'environ 40% des diabétiques sont porteurs d'une rétinopathie, ce qui représenterait environ 1 000 000 de patients en France.[3]

La rétinopathie diabétique est aussi fréquente au cours du diabète de *type 1* qu'au cours du diabète de *type 2* :

- dans le diabète de type 1, la RD ne survient en général pas avant 7 ans d'évolution; après 20 ans d'évolution, 90 à 95 % des diabétiques de type 1 ont une RD, dont 40% une RD proliférante.
- dans le diabète de type 2, 20 % des diabétiques de type 2 ont une RD dès la découverte de leur diabète. Le risque à long terme des diabétiques de type 2 est moins celui d'une rétinopathie proliférante (20%) que celui d'un œdème maculaire (60%). Les principaux facteurs de risque de rétinopathie diabétique sont l'ancienneté du diabète, le mauvais équilibre glycémique et l'hypertension artérielle.

1.4.4 Rétinopathie Diabétique (RD)

Dans ce paragraphe une brève description de la rétinopathie diabétique est présentée.

1.4.4.1 Définition

La rétinopathie diabétique est une maladie dégénérative des vaisseaux sanguins de la rétine, qui survient chez les personnes atteintes de diabète. Elle apparaît lorsque l'excès de sucre dans le sang, caractéristique du diabète, a endommagé les petits vaisseaux qui approvisionnent la rétine en nutriments et en oxygène.

La rétinopathie diabétique (RD) est une maladie à forte prévalence chez les diabétiques. Il existe des traitements pour prévoir sa progression et réduire le risque de cécité. L'efficacité du traitement est plus grande lorsque le diagnostic est posé tôt et qu'il est administré en temps opportun. Après analyse des données probantes, des expériences existant ailleurs et du contexte actuel en Algérie, la mise en place d'un programme de dépistage de la rétinopathie diabétique est pertinente et faisable. Un tel programme vise à déceler les personnes qui présentent une rétinopathie à haut risque, car elles peuvent bénéficier d'un traitement préventif. Avant d'établir des scénarios réalistes et d'estimer leurs coûts, il faudra définir, entre autres, l'organisation du dépistage dans le cadre de la prise en charge globale actuelle des patients diabétiques, le partage des activités de dépistage entre les professionnels, et la formation de professionnels qui devront assurer la qualité des photos [4].

1.4.4.2 Signes Cliniques de la RD

Les différents types de lésions qui peuvent apparaître au cours du développement de la RD sont au nombre de 9, certains d'entre eux se déclinant en sous-catégories. Ils sont listés dans le tableau. 1.1.

Numéro	Type de lésion de RD
1	Micro-anévrismes
2	Hémorragie rétinienne
3	AMIRs (Anomalies Microvasculaires Intrarétiniennes)
4	Anomalies veineuses
5	Ischémie rétinienne
6	Exsudats
7	Nodules cotonneux
8	Néo-vaisseaux

Tableau 1.1 Lésions de la Rétinopathie Diabétique.

1. Micro anévrismes [6]

Les micro-anévrismes sont les premiers signes non équivoques de la RD. Ils sont produits par un affaiblissement local des parois capillaires. Ils apparaissent sous forme de points rouges. Leur taille varie de 10 à 100 microns de diamètre. La figure. I.4 (a) montre un exemple de telles lésions.

2. Hémorragies

La progression de la maladie peut produire une rupture d'une paroi d'un capillaire ou d'un micro anévrisme ce qui provoque des hémorragies rétinienne. Elles apparaissent soit comme des petits points rouges indiscernables à partir de micro-anévrismes ou grandes taches de forme ronde avec contour irrégulier. La figure. I. 4(b) montre certains d'entre eux.

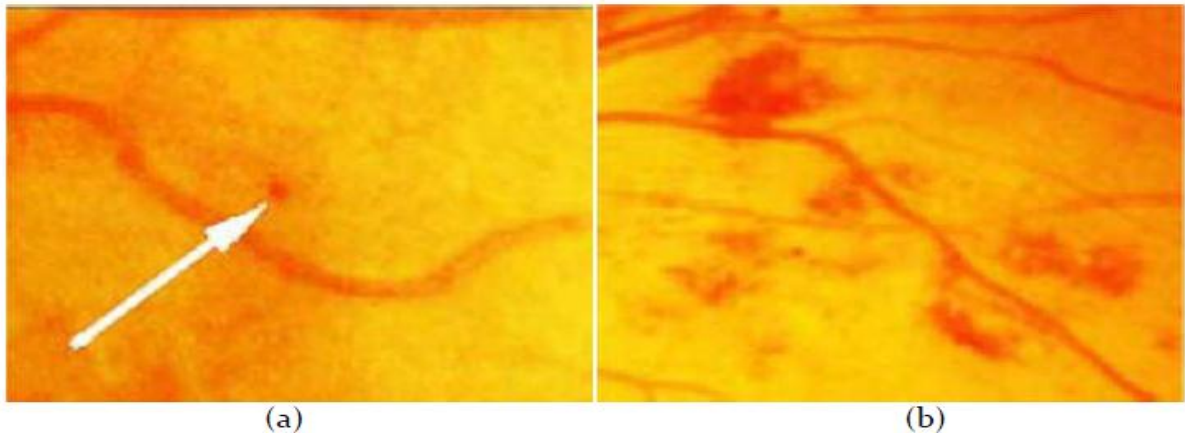


Figure. 1.4 : Présence des micro-anévrismes et hémorragies dans les images couleur de la rétine ; (a) : micro-anévrismes montrés par une flèche ;(b) : hémorragies.

3. Anomalies Microvasculaires Intra Rétiniennes (AMIR)

Sont des anomalies vasculaires comme des dilatations, des ectasies ou des bourgeonnements. Elles témoignent d'une occlusion vasculaire et par conséquent d'une ischémie locale. Elles représentent une réponse neovasculaire à cette occlusion. Leur forme est plutôt irrégulière, mais parfois, elles peuvent être facilement confondues avec des micro-anévrismes (figure .1.5).

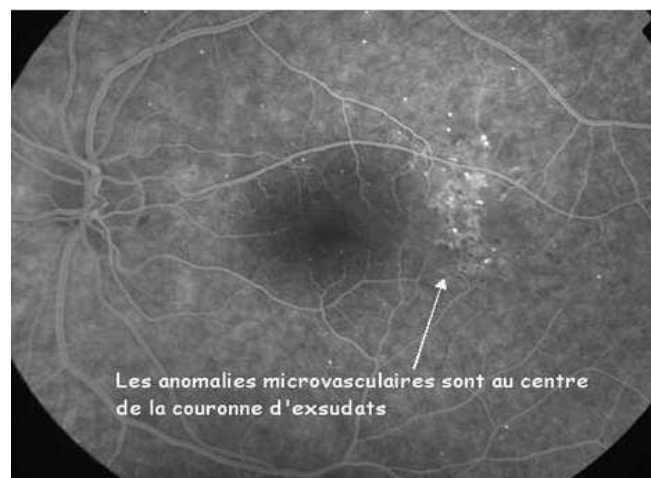


Figure 1.5. Présence des AMIR et une anomalie veineuse rétinienne

1.4.4.3 Stades de la rétinopathie diabétique(en fonction de la sévérité de l'ischémie)

a).rétinienne du risque de néovaisseau pré-rétinien (6)

Voici les stades de la rétinopathie diabétique, en fonction de la sévérité de l'ischémie rétinienne du risque de néovaisseau pré-rétinien :

1. pas de rétinopathie diabétique.
2. rétinopathie diabétique non proliférante
 - RD non proliférante *minime* (micro anévrysmes isolés, hémorragies rétinienne punctiformes peu nombreuses).
 - RD non proliférante *modérée* (nombreux micro anévrysmes, hémorragies rétinienne punctiformes, nodules cotonneux, signes d'ischémie rétinienne peu nombreux)(voir Figure 1.6)

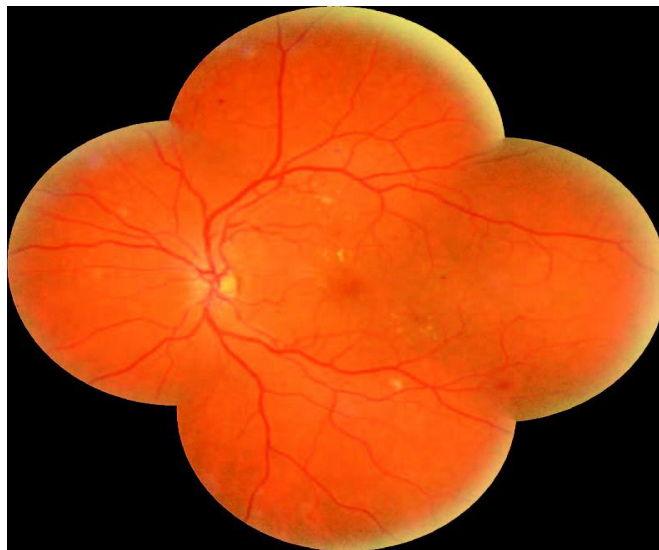


Figure 1.6 RD non proliférante modérée : hémorragie en tache peu nombreuses

- RD non proliférante *sévère* (ou RD *préproliférante* = RDPP), définie par l'association aux signes précédents de signes ophtalmoscopiques évocateurs d'ischémie rétinienne sévère, notamment hémorragies intra-rétiniennes étendues en tache, anomalies veineuses et AMIRs. C'est un stade de haut risque d'évolution vers la néovascularisation, notamment dans certaines circonstances cliniques (cf. IV. Evolution).(voir la figure 1.10)

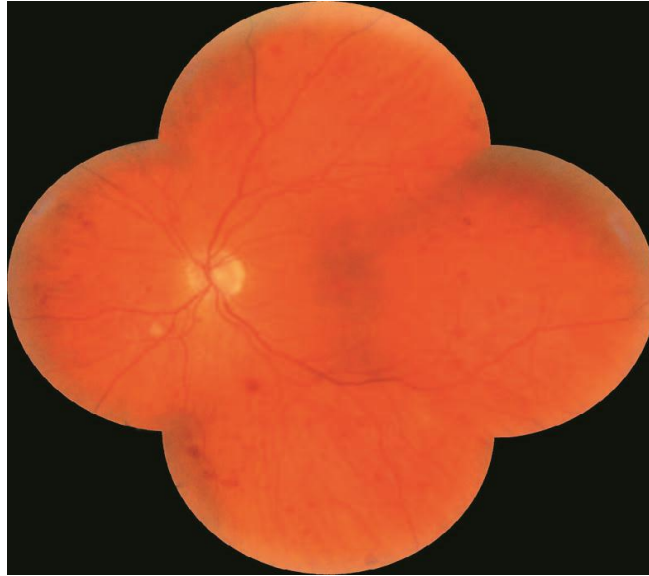


Figure 1.7 RD non proliférante sévère : nombreuses hémorragies en taches sur toute la périphérie

3. rétinopathie diabétique proliférante (RDP)

- RDP *non compliquée* : présence de néovaisseaux pré-rétiniens et/ou prépapillaires.
(voir figure 1.8 ;1.9).

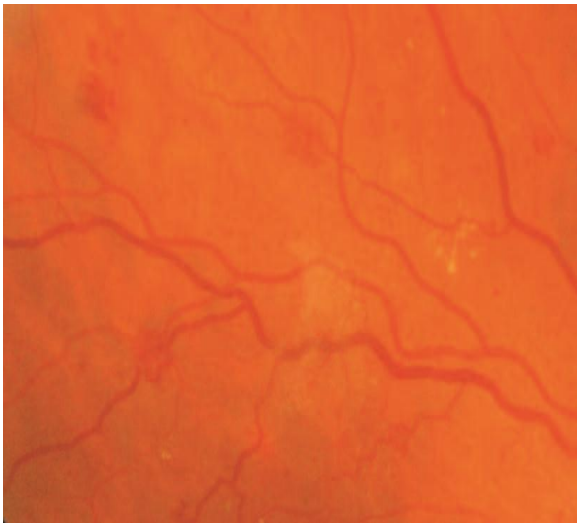


Figure 1.8 Néovaisseau pré-rétinien

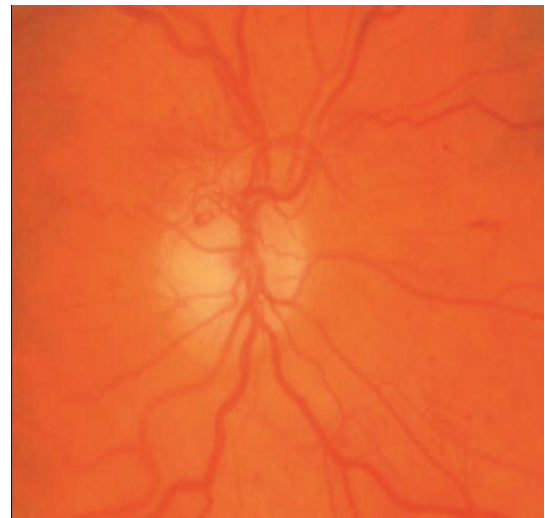


Figure 1.9 Néovaisseaux pré-papillaires

- RDP *compliquée* :
 - hémorragie du vitre,
 - décollement de rétine par traction,
 - glaucome néovasculaire.

Résumé

Pas de rétinopathie diabétique (RD)	+/- maculopathie diabétique
Rétinopathie diabétique non proliférante	
RD non proliférante minime RD non proliférante modérée RD non proliférante sévère (ou RD préproliférante)	
Rétinopathie diabétique proliférante *	
RD proliférante minime RD proliférante modérée RD proliférante sévère RD proliférante compliquée	

*: Proliférante : avec néovaisseaux pré-rétiniens et/ ou pré-papillaires

Tableau 1.2 Classification de la rétinopathie diabétique

b) Maculopathie diabétique

- *œdème maculaire focal* : œdème localisé entouré d'exsudats.(voir figure 1.10)
- *œdème maculaire diffus* : œdème maculaire étendu, lié à une hyperperméabilité généralisée à partir des capillaires du pôle postérieur, non cystoïde ou cystoïde (en fonction de la présence de kystes intra-rétiniens).(voir figure 1.9)
- *maculopathie ischémique*



Figure 1.10 Œdème maculaire focal localisé, entouré d'exsudats

1.4.4.4 Traitement de la rétinopathie diabétique [5]

1. Traitement médical

➤ Équilibre glycémique et tensionnel ++++ :

- l'effet bénéfique chez les diabétiques de type 1 d'une *bonne équilibration glycémique* sur l'incidence et la progression de la rétinopathie diabétique a été démontré par le DCCT (Diabetic Control and Complications Trial Research Group).
- l'effet bénéfique chez les diabétiques de type 2 d'un bon équilibre glycémique mais aussi du *contrôle de l'hypertension artérielle* est également démontré par l'étude UKPDS (United Kingdom Prospective Diabetes Study).

Le contrôle strict de ces 2 facteurs est essentiel pour réduire et ralentir la progression de la RD tant chez les diabétiques de type 1 que de type 2.

➤ Traitements médicamenteux

L'efficacité des antiagrégants plaquettaires est incertaine : leur utilisation, un temps proposée, n'est plus retenue. Une tendance des antagonistes du système rénine angiotensine, notamment du candesartan, a été montrée dans l'étude DIRECT (Diabetic RETinopathy Candesartan Trial) pour réduire l'incidence de la RD chez les type 1 et favoriser la régression des stades précoces de la RD chez les type 2. Néanmoins, l'effet n'a pas été suffisant pour recommander leur indication systématique.

2. Traitement de la rétinopathie diabétique proliférante

➤ Photocoagulation au laser

La photocoagulation panrétinienne (PPR) au laser est le traitement spécifique de la rétinopathie diabétique proliférante. Elle consiste en une coagulation étendue de toute la rétine périphérique : *elle n'agit pas directement sur les néovaisseaux mais indirectement par la destruction étendue des territoires d'ischémie rétinienne* : celle-ci permet d'obtenir la régression de la néovascularisation prérétinienne et/ou prépapillaire dans près de 90% des cas et de réduire considérablement le risque de cécité lie à la rétinopathie diabétique proliférante. Elle est réalisée en ambulatoire sous anesthésie de contact, en plusieurs séances.

La PPR est indiquée :

a) ***dans tous les cas de rétinopathie diabétique proliférante.*** La rapidité de sa réalisation dépendra de la sévérité de la RDP : des neovaisseaux pré-rétiniens isolés et peu nombreux seront traités par des séances espacées d'une ou plusieurs semaines ; à l'autre extrême, l'existence d'une néovascularisation irienne est l'indication d'une PPR urgente, réalisée par des séances très rapprochées.

b) ***dans certains cas de RDNP sévère (RDPP)*** , a titre préventif, chez des sujets à haut risque de passage en rétinopathie proliférante : principalement grossesse, normalisation rapide de la glycémie (notamment par mise sous pompe à insuline), chirurgie de la cataracte. La PPR prévient efficacement dans ces cas la survenue d'une néovascularisation.

➤ Injections intravitréennes d'anti-VEGF

Des injections intravitréennes d'anti-VEGF : bévacizumab (Avastin®), pégaptanib (Macugen®) ou ranibizumab (Lucentis®), comme dans la DMLA, peuvent par leur pouvoir anti-angiogéniques avoir un intérêt dans certaines indications de rétinopathie diabétique proliférante, notamment dans le traitement du glaucome néovasculaire ; il nécessitent cependant des autorisations spéciales, n'ayant pas d'AMM dans le traitement de la rétinopathie diabétique.

➤ Traitement chirurgical (vitrectomie)

Il est indiqué dans les cas de RDP compliquée d'hémorragie intra-vitréenne persistante ou de décollement de rétine tractionnel.

1.5. Conclusion

Avec l'évènement d'outils de traitement des images médicales, des applications comme le dépistage automatique de la rétinopathie diabétique pour la recherche des différents stades pathologiques sont devenues envisageables. La constitution de bases de données d'images apporte une base de connaissances supplémentaires qui permet la création d'ensembles de données de référence utiles à la création de modèles et à l'analyse statistiques d'images.

Dans ce chapitre préliminaire nous avons décrit le cadre ou le contexte biomédical dans lequel s'inscrit notre travail à savoir la Rétinopathie Diabétique, une maladie particulièrement sévère et répandue, et qui affecte la rétine des malades diabétiques entraînant, dans le cas de non dépistage précoce et de survenue de complications, une cécité visuelle.

Nous avons commencé par une description de l'anatomie de l'œil en situant la rétine. Ensuite on a présenté l'exploration de cette dernière (la rétine) à l'aide d'un examen angiographique, et enfin on a détaillée la Rétinopathie Diabétique : le cours de la maladie, ces symptômes et les complications qui peuvent survenir.

Dans le chapitre qui suit, on fera un état de l'art sur la recherche d'images de manière générale : recherche d'images par mots clés, recherche d'images par le contenu visuel dans le cas général et dans le cadre d'imagerie médicale, on abordera aussi les différentes techniques d'extraction de descripteurs visuels à partir d'images numériques.

CHAPITRE 2

RECHERCHE D'IMAGES

2.1. Introduction

L'accroissement constant en nombre et en volume , des bases de données numériques surtout avec l'explosion d'Internet et aussi le développement à grande échelle de l'imagerie numérique, il n'est pas rare d'avoir des bases d'images numériques contenant plusieurs milliers et même plusieurs dizaines de milliers d'images, que ce soit des bases ciblées pour un domaine donné (médecine, journalisme, tourisme, éducation, ...) ou tout simplement pour les particuliers qui accumulent d'immenses bases de photographies numériques (souvenirs, voyages, famille, événements, ...). Pour gérer et utiliser efficacement ces bases d'images, un système d'indexation et de recherche d'images est nécessaire. C'est pourquoi la dernière décennie a vu une explosion du nombre d'études effectuées dans le domaine de l'indexation et de la recherche d'images. Ce sujet devient un sujet très actif, plusieurs équipes de recherche ont développé des travaux dans ce sens, plusieurs secteurs d'activité sont concernés parmi lesquels la géographie, la médecine, les musées, la surveillance et le monitoring, etc. Chaque utilisateur d'un des domaines en question souhaite retrouver les images selon ses propres critères de recherche.

Nous rappelons dans ce chapitre à un état de l'art des différentes approches de représentation et de recherche d'image connus dans la littérature, ensuite les formes de requêtes dans les systèmes de recherche d'images dans chaque approche ainsi leurs limites.

2.2. Domaines d'application de la recherche d'images

La recherche d'images, est en plein essor depuis les années 90 en raison principalement de la diversité des applications. On peut plus particulièrement mettre en avant les domaines où les méthodes de recherche "efficaces" sont nécessaires :

- la médecine pour les dossiers médicaux des patients ;
- Les applications industrielles pour le contrôle de qualité des produits sur les chaînes de fabrication ;
- Le journalisme avec les journaux de presse écrite, les journaux télévisés ou encore les documentaires qui exploitent les bases d'images des agences de presse ;
- L'art et la culture en général, pour les musées ;
- La gestion des catalogues de produits pour les applications de commerce en ligne ;
- Les applications de sécurité ou d'authentification pour l'identification de visages ou d'empreintes dans le domaine de la biométrie.
- À ces formes d'utilisation plus ou moins spécialisées, on peut ajouter un fort développement de l'utilisation des images numériques par et pour des utilisateurs non spécialistes (scanner, appareil photo numérique, etc.) qui contribuent au développement de bases d'images généralistes.
- L'internet : c'est la recherche d'image sur le WEB.

2. 3. Bases d'images

On distingue généralement deux types de bases d'images pour lesquelles les processus de recherche sont appliqués :

- les bases d'images appartenant à un domaine spécifique : (ex : images médical, images satellites, bases d'empreintes, bases de visages, etc.). ce domaine se caractérise par une faible variabilité inter-images (généralement même conditions de prise de vue, les objets sont connus à priori, etc.) et il présente donc une forte corrélation entre le contenu numérique et les concepts sémantiques correspondants.
- Les bases d'images généralistes : les images sont de nature différente, hétérogènes, des caractéristiques différentes, et leurs sémantiques n'est décrite que partiellement. (Différentes conditions de prises de vues, les objets ne sont pas connus à priori, chaque objet peut être interprété de plusieurs façons possibles, etc.).

2.4. Approches de représentation et de recherche d'images

Il existe trois approches de représentation et de recherche d'images connues dans la littérature [9] à savoir:

- **L'approche orientée-contexte** : considère l'image comme une boîte noire. Il s'intéresse à l'environnement global de l'image : son auteur, sa date d'acquisition, etc. Par exemple, « cette image a été prise par le Dr Robert en juin 99 » est une description admise par cette approche.
- **L'approche orientée-contenu** : utilise *la description objective* pour représenter des images. Il s'intéresse à la manipulation du contenu physique de l'image par le biais de ses caractéristiques physiques comme la couleur, la texture, la brillance, les formes, etc. « Rechercher les images contenant une couleur donnée et une brillance spécifique à telle position » est un exemple typique des requêtes de cette approche.
- **L'approche orientée-sémantique** : se focalise sur l'interprétation de l'image en utilisant *la description subjective*. Il décrit les objets qui composent l'image et leurs relations comme ils sont définis et perçus par l'être humain. « Rechercher toutes les images contenant une tumeur » est un exemple de requête orientée sémantique.

Les systèmes actuels sont élaborés en fonction des besoins liés à chaque domaine d'application. Par exemple, les applications de surveillance se basent sur les approches orientée-contenu et orientée-contexte selon lesquels la recherche se fonde, d'une part, sur la date d'acquisition de l'image, et, d'autre part, sur ses formes, ses couleurs et ses textures. Dans le domaine médical les modes de recherche sont basés sur le contenu ainsi que la sémantique de l'image et ça se qui nous intéresse (l'imagerie médicale) dans notre projet.

2.4.1 Approche orientée-contexte

2.4.1.1 Définition

L'approche orientée-contexte regroupe l'ensemble des systèmes qui travaillent sur le contexte de l'image et non plus sur son contenu (le contenu de l'image est une boîte noire). Il utilise le contexte comme moyen de description, de recherche et d'indexation d'image.

Par définition, le contexte d'une image est l'ensemble des informations attachées à cette image sous forme d'une annotation textuelle, permettant ainsi d'orienter sa signification, la photo de la figure 2.1 peut être interprétée comme étant celle d'une montagne si on ignore le texte qui l'accompagne Il s'agit en réalité du "Lac du Passet".



Figure 2.1 Lac du Passet ([9])

2.4.1.2 Indexation du contexte de l'image

Le contexte se compose d'un ensemble d'informations factuelles (annotations textuelles). Or, l'indexation des informations textuelles est étudiée depuis longtemps et les méthodes élaborées sont nombreuses. Tel que l'indexation plein texte (Full Text Scanning), la signature ([9]).

La description textuelle est une opération longue, coûteuse et difficile à élaborer car l'information externe est manuellement attachée par l'utilisateur ce qui conditionne la qualité de recherche future, et puis elle ne décrit pas fidèlement le contenu de l'image car elle se fait de manière automatique à partir du nom, de la légende ou du texte qui l'entoure.

2.4.1.3 Formes de requêtes dans les systèmes de recherche d'images par le contexte[21]

Les systèmes de recherche d'image par le contexte travaillent en général avec des requêtes par *mots clés*. Afin de faire ce travail le système se base sur l'annotation textuelle d'images. Beaucoup de moteurs de recherche d'images tels que Google, Yahoo utilisent cette façon. Mais elle n'est pas parfaite parce que quelques mots n'expriment pas le sens d'une image. La figure 2.2 donne un exemple, l'utilisateur veut trouver des images contiennent une (des) voiture(s) avec le ciel. Cependant les premières images apparaissant ne sont pas pertinentes car ils n'ont ni les voitures ni le ciel.

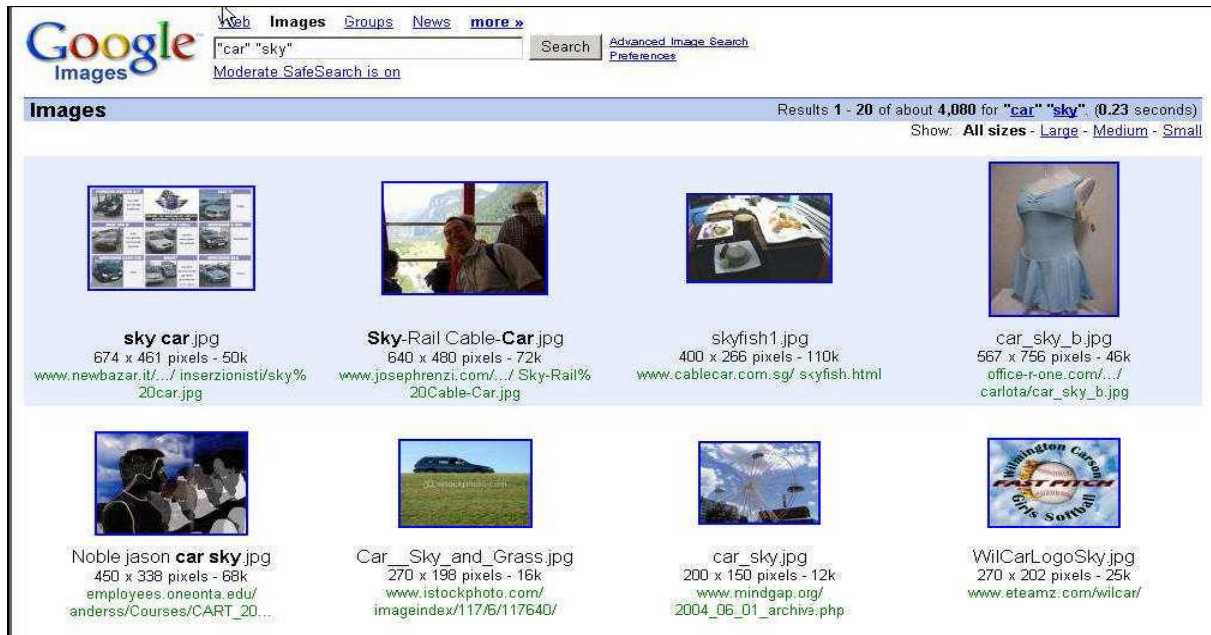


Figure 2.2 Exemple de recherche d'images dans Google.

2.4.1.4 Quelques Systèmes de l'approche orientée-contexte

La première génération de systèmes proposés de gestion de base d'images a adopté l'approche orientée-contexte pour une raison évidente : les techniques de description et de gestion des informations textuelles sont étudiées depuis longtemps par la communauté scientifique. Ces systèmes ne sont pas nombreux. Le contexte est souvent adopté par les moteurs de recherche d'images sur le WEB.

- **PACS (Picture Archival and Communication Systems)** systèmes de communication et d'archivage d'images. Ce sont des systèmes de gestion d'image qui ont adopté ce paradigme
- **I.S.I.S. (Interactive System for Image Sélection)** est un système qui permet à l'utilisateur d'accéder à une base d'images médicales. Il fait partie d'un système d'information hospitalier (HIS) DIOGENE ([36]).
- **Système AMORE (Advanced Multimedia Oriented Retrieval Engine)** ([9]) est un moteur de recherche d'images sur le WEB qui a adopté les paradigmes orienté-contexte et orienté-contenu. Il permet de rechercher des images grâce à des mots-clés, des thèmes, et des comparaisons d'images. AMORE utilise plusieurs composants pour rechercher les images : Glimpse, Harvest Web Indexing , WordNet.

2.4.1.5 Limites de l'approche orienté-contexte

la description grâce au contexte présente certaines limites et lacunes:

- Les besoins des utilisateurs dans certains domaines, en particulier le domaine médical, nécessitent des modes de recherche différents basés sur le contenu de l'image. Par exemple, dans un traitement thérapeutique, le médecin est parfois amené à vouloir retrouver des cas en faisant intervenir certains attributs d'image tels que : la densité des objets, la texture, la position, la forme, etc. Ces attributs permettent au médecin de comparer les images et l'aident à établir son diagnostic. Dans ce cas, le contexte seul est incapable de décrire les images de façon à satisfaire les requêtes des utilisateurs concernés. Cependant, il doit être intégré dans la représentation de l'image médicale d'une part pour répondre aux requêtes basées sur les informations contextuelles et, d'autre part, pour compléter la description de l'image.
- Le contexte n'existe pas nécessairement dans certains domaines. Ainsi, les images filmées par une caméra de surveillance, par exemple, ne contiennent pas d'informations contextuelles capables de décrire l'image. Les seules informations connues sont la date et le lieu de la scène. Par conséquent, l'utilisation d'un tel paradigme dans ce domaine d'application peut s'avérer inadéquate pour décrire l'image.
- Les méthodes de recherche d'images orientées-contexte sont limitées aux langages textuels et à l'hypermédia. Or, ceux-ci ne peuvent pas exprimer précisément et facilement tous les besoins de l'utilisateur.
- La recherche d'images par mots clés pose le Problème de *Bruit et Silence*. Bruit : des images non pertinentes sont retournées par la recherche parce qu'elles contiennent les mêmes mots clés mais dans un autre contexte, Silence : des images pertinentes non retournées par la recherche parce qu'elles contiennent des concepts synonymes aux mots clés.

Les besoins des utilisateurs et l'évolution de la recherche d'images ont conduit à un constat : les systèmes orientés-contexte ou contextuels sont très restrictifs et inappropriés. Pour pallier aux inconvénients de la recherche par mots clés, une deuxième approche a été proposée : la recherche par le contenu. Nous verrons par la suite de ce chapitre que la description de l'image a pris deux directions : la description physique et la description sémantique de son contenu.

2.4.2 Approche orienté-contenu

2.4.2.1 Définition

L'approche orienté-contenu regroupe l'ensemble des systèmes qui proposent de décrire et de rechercher les images par le biais de ses caractéristiques physiques. Ces caractéristiques, encore appelées caractéristiques de *bas-niveau*, symboliques, ou bien visuelles sont: la distribution des couleurs (les histogrammes), la texture, la forme, la brillance,...etc.

L'utilisation de cette approche a permis une nouvelle souplesse et une indépendance certaine vis-à-vis du type de l'image et du domaine d'application.

La recherche dans tels systèmes se fait par le calcul de distance de similarité entre deux images. Le système prend une ou des requêtes à l'utilisateur et lui donne le résultat correspond à une liste d'images ordonnées en fonction de la similarité entre ces images et l'image requête [21].

2.4.2.2 Caractéristiques physiques de l'image et la recherche

Les recherches basées sur le contenu physique de l'image peuvent porter sur ses caractéristiques physiques globales et locales.

➤ **Les caractéristiques physiques globales**

Concernent l'image entière, et sont calculés de façon globale sur l'image. Parmi ces caractéristiques : la couleur, la texture et la forme.

L'utilisation de ces caractéristiques offre différents avantages [9]:

- L'interrogation des caractéristiques globales de l'image est mieux adaptée aux requêtes qui touchent à la globalité de l'image. "Trouver les images dont le rouge est la couleur dominante" est un exemple type de requête basée sur les caractéristiques physiques globales de l'image.
- Les caractéristiques physiques globales de l'image sont utilisées lors de la détection des images (ou frames) dans une vidéo afin d'identifier les changements de scènes.
- L'extraction des caractéristiques physiques globales nécessite un temps de calcul faible.

Cependant, L'inconvénient majeur de cette utilisation concerne le nombre important d'images non pertinentes obtenu lors d'une requête basée sur des caractéristiques physiques globales. Un système basé uniquement sur des caractéristiques globales ne peut pas donner les résultats désirés, surtout dans le cas où l'image est composée de plusieurs objets ayant des caractéristiques, couleurs et textures, très différentes.

➤ Les caractéristiques physiques locales

Concernent les objets de l'image. Ils portent sur la forme des objets, leurs bordures, la disposition des couleurs, ... etc.

Pour calculer les caractéristiques locales deux approches différentes peuvent être employées. La première approche consiste à diviser une image en utilisant une grille et les caractéristiques sont calculées pour chaque case de cette grille. Par exemple dans la figure 2.3(a) l'image est divisée en 12 zones de même taille et les caractéristiques sont calculées pour chacune de ces zones. La deuxième approche consiste à segmenter l'image pour la diviser en zones locales plus proches des objets constituant l'image et ensuite calculer les caractéristiques pour chacune des régions extraites (figure 2.3(b)). Pour calculer ces caractéristiques, différentes techniques sont utilisées telle que : la distribution des couleurs, les histogrammes des couleurs, la texture, la transformée de Fourier, ... etc.



Figure 2.3 Deux découpages différents de l'image.

(a) Découpage en grille de l'image, (b) Découpage par la segmentation

Le choix des caractéristiques locales dans une recherche dépend du domaine d'application. Par exemple, dans une radiographie, nous nous intéressons aux formes, positions et dimensions des objets dans l'image, tandis que dans une IRM, les couleurs sont plus importantes.

Malgré l'utilité de rechercher des images en se basant sur ses caractéristiques locales, leur utilisation seule peut présenter l'inconvénient de perte de sens global de l'image, en submergeant celui-ci dans un flot de petits détails inutiles. Par conséquent, un compromis doit être trouvé, selon les applications et selon les requêtes individuelles, entre caractéristiques globales et caractéristiques locales. [9]

2.4.2.3 Indexation des caractéristiques physiques de l'image

L'indexation des caractéristiques physiques de l'image demeure un domaine de recherche assez ouvert. Plusieurs méthodes d'indexation du contenu physique ont été proposées. Un survol des méthodes d'indexation spatiales existe dans [32].

Dans [9] une classification proposée est comme suit :

- Les méthodes qui transforment les rectangles en points dans un espace à n dimensions.
- Les méthodes de Z-ordering [29],
- Les méthodes de comparaison des vecteurs de caractéristiques ,
- Les méthodes de représentation par arbre : sont les plus utilisées. Elles utilisent une représentation multidimensionnelle, qui consiste à découper l'image en plusieurs niveaux en suivant certains critères d'une manière récursive. On trouve la Méthode de représentation d'arbre R-tree [15], Quadtree [14], k-d tree [7],...etc.

2.4.2.4 Formes de requêtes dans l'approche orienté-contenu [21]

Les systèmes recherche d'images par le contenu utilisent souvent deux formes de requêtes: requêtes par esquisse et requêtes par image exemple.

➤ Requêtes par esquisse

Dans ce cas, le système fournit à l'utilisateur des outils lui permettant de constituer une esquisse (dessin) correspondant à ses besoins. L'esquisse fournie sera utilisé comme exemple pour la recherche. L'esquisse peut être une ébauche de forme ou contour d'une image entière (caractéristiques physiques globales) ou une ébauche des couleurs ou textures des régions d'une image (caractéristiques physiques locales). L'utilisateur choisira, en fonction de la base d'images utilisée, de ses besoins et préférences, l'une ou l'autre de ces représentations. Le choix des caractéristiques dépend du système et surtout du domaine d'application visé.

Cette technique présente l'inconvénient majeur qu'il est parfois difficile pour l'utilisateur de fournir une esquisse, malgré les outils qui lui sont fournis.

Parmi les caractéristiques les plus utilisées:

- la couleur

La couleur [24] est un des premiers descripteurs employés pour la recherche d'images, La couleur forme une partie significative de vision humaine, sans elle beaucoup de tâches

journalières prouveraient très difficile. Nous pouvons distinguer efficacement les objets basés sur seule la couleur.

Il existe plusieurs modes de représentation de la couleur : les espaces des couleurs (le modèle *RGB*, *Lab*, *XYZ* et *HSV* (Hue-Saturation-Value) (aussi connu sous le nom de système de cône hexagonal)), l'histogramme, les moments statistiques, la cohérence spatiale, les couleurs dominantes, le corrélogramme.

L'information relative aux couleurs est particulièrement importante dans la caractérisation d'une image. Plusieurs études ont été menées pour trouver un critère de choix des descripteurs de couleurs pour l'indexation des images, mais aucune n'a abouti. Ceci peut s'expliquer par le manque de subjectivité de cette information, les descripteurs couleur ne suffisent pas à indexer efficacement une image, ni à la chercher. [16]

- la forme

La recherche à base de forme consiste à comparer deux images ayant des objets de même forme. On propose à l'utilisateur de formuler sa requête en utilisant des formes prédéfinies à comparer à l'ensemble des images stockées dans la base. Les recherches basées sur la forme donnent des résultats assez satisfaisants. De nombreuses méthodes sont référencées dans la littérature [13] pour l'analyse de formes. Parmi les plus connues, on peut citer : Les descripteurs géométriques de région [[22], [[23]]], Les moments géométriques [20], Transformée de Hough [18]. La description des formes est une tâche difficile. Ainsi dans les systèmes de recherche récents, peu emploient la forme comme critère de recherche [21]. La figure 2.4 montre un exemple dans le système **QBIC**. Ce système permet de créer une requête avec les formes et les couleurs des objets.



Figure 2.4 Une requête par esquisse dans QBIC

- la texture

La texture, autre primitive visuelle, est étudiée depuis une vingtaine d'années et plusieurs techniques ont été développées pour l'analyser tel que la matrice de cooccurrences de Haralick et les méthodes basées sur les filtres de Gabor [21].

Dans quelques cas, il est difficile de faire l'esquisse et il est aussi difficile de détecter les différences avec les caractéristiques de l'image. Donc en général, le système utilise seulement le paradigme requête par image exemple ou utilise les deux approches ensembles.

➤ Requêtes par image exemple

Dans ce cas le système a besoin de comparer un exemple de même type (image) avec la base pour produire les documents similaires. Cette méthode est simple naturelle et ne nécessite pas de connaissance approfondies pour manipuler le système. Elle est donc bien adaptée à un utilisateur non spécialiste.



Figure 2.5 Une requête par image dans Blobworld

Par exemple, dans le système Blobworld, pour choisir une requête, le système affiche quelques classes d'images tels que 'Animal', 'OceanScenes', 'Outdoor Scenes', 'People', 'ManmadeObjects'. L'utilisateur clique sur une classe pour voir quelques images dans cette classe et choisit une de ces images comme image requête. Dans la figure 2.5 l'utilisateur clique sur la classe 'Flower' et quelques images de fleurs sont affichées.

Un autre exemple avec le système CBIR (Content Based Image Retrieval) dans le domaine médical, qui sert comme aide aux diagnostique pour les médecins. Par l'introduction d'une image requête, le système retourne les images de diagnostique les plus similaires avec les cas de traitement correspondants.

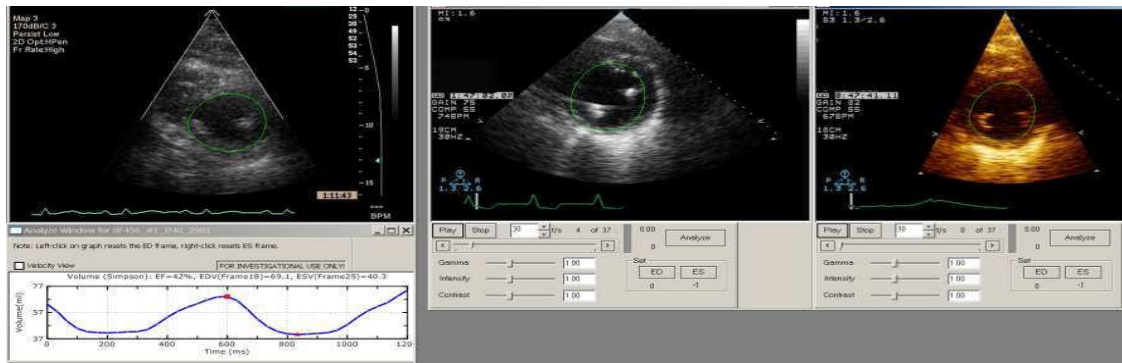


Figure 2.6 Requête par image exemple dans le système CBIR [37]

L'utilisation des requêtes par exemples est très visuelle et très souvent utilisée. Mais cette approche a un inconvénient. Les systèmes supposent que plus il y a d'exemples disponibles, meilleurs sont les résultats. Mais, les exemples additionnels peuvent contenir des caractéristiques inappropriées et faire baisser la performance du système de recherche [8].

2.4.2.5 Quelques Systèmes de l'approche orienté-contenu

Le premier prototype de système a été proposé en 1970 et ce système a attiré l'attention de beaucoup de chercheurs [9]. Maintenant, plusieurs systèmes de recherche d'images sont disponibles sur le marché. Certains sont commercialisés avec une démonstration sur le WEB, d'autres restent en phase expérimentale. Nous avons choisi certains systèmes et prototypes parmi les plus connus :

➤ **Le système Q.B.I.C. (Query By Image Content)** développé dans le centre de recherche IBM d'Almaden [28], est considéré comme le système de recherche d'images le plus connu. Il gère non seulement une collection d'images hétérogènes, mais aussi des vidéos. Q.B.I.C. existe en plusieurs versions: démonstration (<http://wwwqbic.almaden.ibm.com/>), application ou bien partie intégrante des autres produits IBM tel que DB2 Digital Library. Dans Q.B.I.C., la description du contenu de l'image est automatique, elle est réalisée par le biais des procédures de reconnaissance de formes, d'extraction des contours des objets de l'image, des textures et de l'histogramme de couleurs. Q.B.I.C. propose également de décrire le contenu de l'image par le biais de mots-clés.

Le processus d'évaluation des requêtes calcule les mesures de similarité (en pourcentage) entre les différentes caractéristiques (croquis, couleur, texture, etc.) de l'image telle qu'elle est décrite par la requête et les caractéristiques physiques des images de la base. Pour une base de 1000 images, le temps de réponse à une requête est compris entre 2 et 40 secondes environ.

Le résultat d'une requête est restitué sous la forme d'une liste d'images classées par ordre décroissant de pertinence.

- **Le système VIRAGE** développé par la société Virage Incorporation, est un système commercialisé connu, présenté sous la forme d'un ensemble de modules. Virage peut être intégré à n'importe quelle application. En particulier, Oracle, dans ses deux versions 8i et 9i, intègre le système Virage pour rechercher l'image par ses caractéristiques physiques (couleur, structure et texture). Virage existe en version d'évaluation (<http://www.virage.com/online/>). Il a été adopté par Altavista's AV Photo Finder (<http://image.altavista.com/>) pour permettre la recherche d'images par son moteur de recherche.
- **Le système Excalibur** développé par la société Excalibur Technologies, offre une variété de techniques d'indexation et de correspondance basées sur la reconnaissance de formes. Une démonstration existe sur le Web (<http://www.excalib.com/>). Le catalogue et le moteur de recherche Yahoo a intégré Excalibur dans Yahoo Image Surfer (<http://isurf.yahoo.com/>). Ce dernier permet ainsi une recherche d'images sur le WEB.

Une grande panoplie de systèmes expérimentaux ont été développés dans l'objectif de prouver la faisabilité des nouvelles techniques de recherche d'images.

- **Photobook** [Hyv, Sty][30] propose une interrogation par le biais de la couleur, la texture, la forme, et autres caractéristiques physiques. Une démonstration en ligne est disponible sur <http://wwwwhite.media.mit.edu/vismod/demos/photobook/>.
- **VisualSeek** [Min][34]. Application développée à l'université de Colombie, elle permet une interrogation basée sur la couleur, la forme, l'aspect spatial, et les mots-clés.
- **WebSeek** [Yam, Oh][35] est un système qui permet, par le biais d'agents, de rechercher des images sur le WEB à travers les caractéristiques physiques. Les images sont également indexées par mots-clé. Une démonstration en ligne existe sur <http://disney.ctr.columbia.edu/WebSEEK/>.

Dans le domaine médical, plusieurs logiciels standalones ont été proposés. Citons le logiciel ANALYZER [31], développé par la Mayo Clinic, qui propose une variété d'opérations de traitement d'images. Un autre logiciel, MEDx [33], permet d'analyser les IRM.

2.4.2.6 Limites de l'approche orienté-contenu

La description des caractéristiques physiques de l'image offre certains avantages : elle est souvent automatisable, et permet l'interrogation graphique. Les résultats obtenus sont assez satisfaisants. Le grand avantage de ces méthodes réside dans leurs bonnes performances.

Mais malgré le réel succès actuel de l'approche orienté-contenu, le problème de la recherche d'images reste ouvert. La recherche des images par les caractéristiques physiques demeure en effet inappropriée dans plusieurs domaines. La description du contenu physique de l'image se heurte ainsi à plusieurs problèmes dans le domaine médical :

- La variété et la complexité du corps humain: Par exemple, la recherche d'une forme circulaire dans une base d'images médicales pourrait correspondre aux radiographies cérébrales ou à des images de cellules.
- La pertinence: La recherche d'une image médicale qui se baserait uniquement sur ses caractéristiques physiques pourrait s'avérer inappropriée. Par exemple, deux échographies contenant les mêmes objets mais prises chacune avec une rotation différente, ont de grandes chances d'être jugées non similaires par un système basé sur un calcul de similarité entre caractéristiques physiques. Les composants de l'image médicale sont plus importants et plus significatifs pour l'utilisateur final que l'image elle-même. Dans ce cas, il est recommandé d'utiliser une méthode de recherche indépendante des supports physiques, de la rotation, du zoom, etc. afin de garantir une cohérence et une efficacité maximales.
- Utiliser seules ces caractéristiques symboliques ne peut pas satisfaire pleinement les requêtes des utilisateurs. La raison principale de cet état de fait est qu'un utilisateur pense sa requête en termes sémantiques (une montagne, une maison ou une personne connue), et non en termes de valeurs symboliques extraites (zone rouge, texture rayée, etc.) des images.
- La multiplicité des images médicales : Le domaine médical fait partie des domaines dans lesquels plusieurs techniques d'acquisition sont utilisées (Scanner, IRM,

échographie, etc.). Or, chacune nécessite un type particulier de traitement d'images et de descripteurs spécifiques.

- Les utilisateurs recherchent régulièrement des concepts (des anomalies) indépendamment du support d'imagerie utilisé. Or, ces méthodes ne proposent aucune compréhension de l'image ni de son contenu tel qu'il est perçu par l'utilisateur.

Donc, un système de recherche d'images par le contenu ne satisfait pas toujours les besoins des utilisateurs. Par conséquent, l'addition de valeurs sémantiques dans tels systèmes est nécessaire.

2.4.3 Approche orienté-sémantique

2.4.3.1 Définition

L'approche orienté-sémantique tente de donner un sens à l'image telle qu'elle est perçue par l'être humain. Il correspond à la description subjective de l'image traduite par des moyens textuels. La figure suivante montre un exemple de description d'une image par la sémantique.

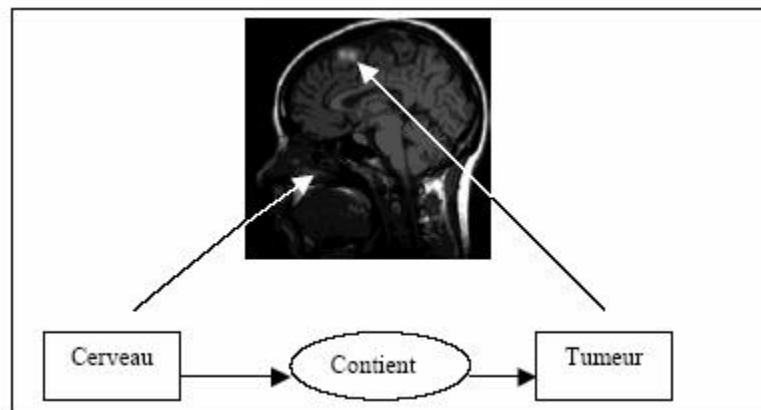


Figure 2.7 Exemple d'une description de radiographie cérébrale par le contenu sémantique[9]

Les approches de description du contenu sémantique de l'image sont nombreuses. Elles sont basées sur la perception et l'interprétation humaine. Nous pouvons les classées en deux Catégories :

- *la première catégorie* : Approches basées sur les travaux effectués dans les deux approches précédentes, et on ajoutant une surcouche de la sémantique. Nous pouvons tirer deux méthodes : *les méthodes textuelles et les méthodes visuelles*.

- *la deuxième catégorie* : Avec les avancées des recherches sur l'aspect sémantique pour trouver des images pertinentes, d'autres approches plus efficaces sont introduites pour décrire

l'aspect sémantique des images : approche basée sur l'*interaction avec l'utilisateur (le retour de pertinence)* et approche basée sur la définition des concepts de l'image par une/des *ontologies*.

Dans ce qui suit nous allons présenter chacune de ces approches.

2.4.3.2 Approches pour la description et l'indexation du contenu sémantique de l'image

2.4.3.2.1 Méthodes textuelles

Permettent une description de l'image par des moyens textuels. Une méthode la plus commune est celle par mots clés [9]. Le principe de cette méthode est le suivant :

- Pour décrire un domaine, souvent un expert dans ce domaine définit l'ensemble des mots clés qu'il juge pertinent. Pour cela il se base sur un corpus de domaine.
- une image dans ce domaine sera donc représentée (indexée) par un sous-ensemble de ces mots clés et qui décrivent son contenu. Des exemples de méthodes d'indexation sont présentés dans [21].

Dans le domaine de la recherche d'images médicales par mots-clé, nous citons comme exemple le site WEB du collège français d'échographie foetale (<http://www.cfef.org/echotheque2/echounefr.php3>) et celui de la bibliothèque internationale francophone d'imagerie médicale (<http://salines.cict.fr:8890/>). Ces sites proposent de formuler les requêtes par un seul mot-clé, par deux mots-clé ou par une image thématique.

2.4.3.2.2 Méthodes visuelles

Ces méthodes utilisent les caractéristiques physiques de l'image (forme, couleur, texture,...etc.) comme base pour décrire son contenu sémantique. Pour cela des représentations visuelles comme des icônes, des images de comparaison, des outils de dessin,...etc. sont associés à des concepts dans le domaine (ceci la différence avec les méthodes orientées contexte) pour décrire une image.

Donc par une simple action (déplacement d'une icône, dessin d'une forme, choix d'une couleur, ...etc.) de la part de l'utilisateur, le contenu sémantique de l'image est décrit et transformé en un ensemble de descripteurs textuels.

Avec ce principe, nous pouvons constater clairement la différence avec les méthodes d'interrogations (formes de requêtes) de l'approche orientée physique. Les méthodes visuelles permettent d'ajouter l'aspect sémantique dans la recherche.

En regardant les avancées récentes en matière de recherche, deux autres approches sont possibles pour ajouter de la sémantique dans un processus d'indexation et de recherche d'images :

- (1) approche basée sur l'interaction avec l'utilisateur (le retour de pertinence) pour comprendre son objectif et le sens de sa requête.
- (2) approche basée sur la définition des concepts de l'image par une/des ontologie(s) ou d'autres mécanismes de représentation de connaissances pour « connecter » (ou lier) la connaissance sémantique humaine avec l'apparence de l'image (les caractéristiques extraites de celle-ci). Ces concepts reflètent la connaissance sémantique humaine.

Une ontologie permet de définir la structure des concepts présents dans l'image, tandis que l'interaction permet de définir les intentions de l'utilisateur dans une requête.

2.4.3.2.3 Retour de pertinence

Pour pallier le manque de connaissances, une approche utilisée dans la recherche d'images consiste à concevoir des systèmes semi-automatiques. On suppose que l'humain maîtrise la sémantique et qu'avec l'aide de quelques interactions simples et ciblées (par exemple, l'identification d'exemples positifs et négatifs à partir d'un premier retour du programme), le système apprend les intentions de l'utilisateur, peut prendre cette sémantique et rend des résultats qui peuvent satisfaire l'utilisateur. Ce mécanisme est appelé **le retour de pertinence** (relevance feedback en anglais).

Dans ce cas-ci, la sémantique n'est pas dans le système informatique, mais elle est extraite implicitement à chaque coup via l'interaction et l'analyse des retours. Le processus de retour de pertinence peut se décrire comme suit :

Quand une requête est faite et qu'une réponse est donnée par le système, l'utilisateur donne plus d'indications au système comme des images *positives* et des images *négatives* à partir des premiers résultats. Le système recalcule et donne de nouveaux les deuxièmes résultats. L'utilisateur peut répéter ce processus plusieurs fois jusqu'à ce qu'il soit satisfait. Les images positives sont les images que l'utilisateur considère comme les images qui sont semblables à la requête et inversement pour les images négatives.

La figure 2.8 donne un exemple de retour de pertinence dans un système de recherche d'images [11]. Dans cet exemple, le processus de retour de pertinence est répété deux fois, le nombre d'images retrouvées pertinentes a augmenté.

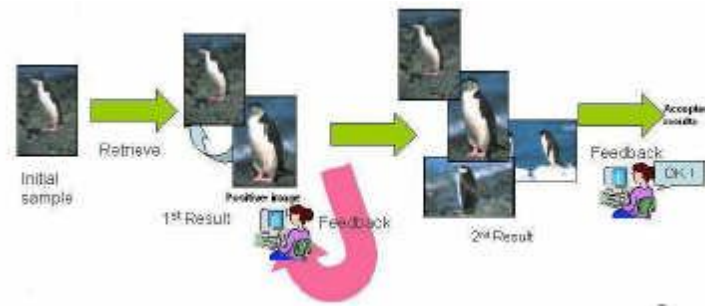


Figure 2.8 Un exemple de retour de pertinence

Le retour de pertinence a des avantages. Cependant, l'utilisateur doit cliquer plusieurs fois pour choisir des images positives et négatives. Ceci n'est pas confortable pour lui.

[21] a proposé une autre interface qui peut faciliter et diminuer les opérations de l'utilisateur. Mais malgré cela, [21] juge qu'avec seulement les mécanismes de retour de pertinence, l'utilisateur ne peut pas être satisfait, il manque un mécanisme robuste. La solution donc est de les intégrer avec d'autres mécanismes.

4.3.2.4 Description des concepts de l'image

La description des concepts de l'image est la définition de termes sémantiques que nous voulons utiliser pour décrire cette image. Cela peut être faite grâce aux ontologies (ou autres mécanismes de représentation de connaissances).

Les informations contenues dans une ontologie peuvent être utilisées pour répondre à des requêtes de l'utilisateur. Pour cela, les images sont annotées par des concepts de l'ontologie et la recherche d'image se fait ainsi en se basant sur ces annotations sémantiques.

Il existe deux approches de recherche d'images se basant sur l'utilisation des ontologies. [21].

□ **Approche utilise une ontologie basée sur les annotations:** consiste à construire une ontologie basée sur les annotations et l'utiliser pour rechercher des images [19]. Cette approche donne des bons résultats puisqu'elle se base sur les annotations (recherche textuelle sans utiliser l'image). Mais elle a des inconvénients puisqu'elle demande beaucoup d'efforts pour faire cette annotation.

Dans [19], les auteurs [Lafon, Keller, Ronald] ont créé une ontologie pour la base d'images du Helsinki University Museum. Ils l'ont utilisée pour la phase d'annotation ainsi que pour la phase de recherche d'images. Les auteurs ont constaté deux problèmes avec leur ontologie.

Quand on fait les annotations, cela peut provoquer des changements dans l'ontologie. Donc comment gère-t-on ces changements pour ne pas demander à l'annotateur de refaire les annotations ? Quand les programmeurs demandent l'ontologie et les annotations de nouvelles demandes pour satisfaire l'utilisateur, l'ontologie ainsi que les annotations doivent être changées.

□ **Approche utilise une ontologie de concepts:** Ces concepts peuvent concerner les données numériques qui sont extraites de l'image et qui concernent ses caractéristiques physiques [26]. Dans ce cas l'approche est semblable aux méthodes visuelles évoquées précédemment mais ici on utilise une ontologie de concepts visuels. Ou bien, ces concepts peuvent décrire le contenu sémantique des images.

Dans [25] les auteurs [Smith, Chang] ont créé une ontologie de concepts visuels. Cette ontologie contient 103 concepts visuels (concepts spatiaux, concepts couleur, concepts texture). Ces concepts sont représentés par un arbre de concepts visuels. Les auteurs ont essayé de l'appliquer au système de recherche d'images. Une base contenant 2000 images a été utilisée pour tester. Les résultats obtenus sont très bons.

La figure 2.9 représente l'interface d'un système d'annotation d'images réalisé par [Landre][15]. Les concepts (mots clés) sont affichés sous forme d'arbre pour plus de facilité.

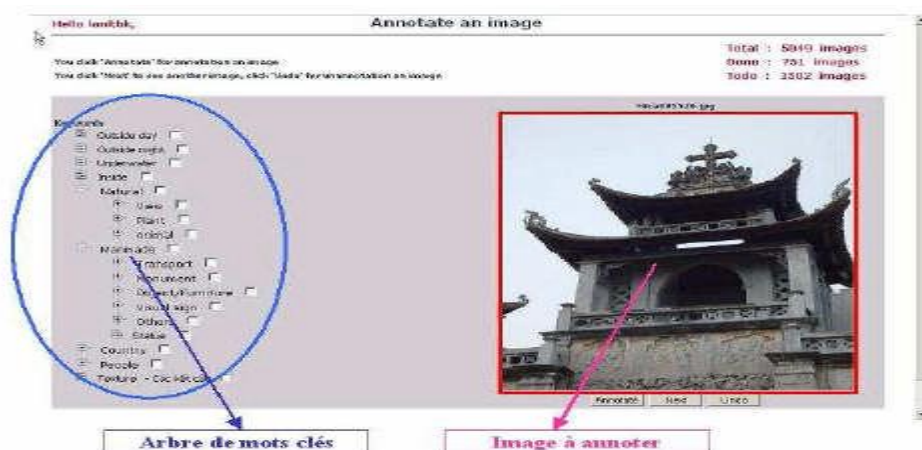


Figure 2.9 Interface pour annoter une image,
Un arbre de mots clés (à gauche), image à annoter (à droite) [21].

La figure 2.10 représente quelques images avec leurs mots clés choisis.

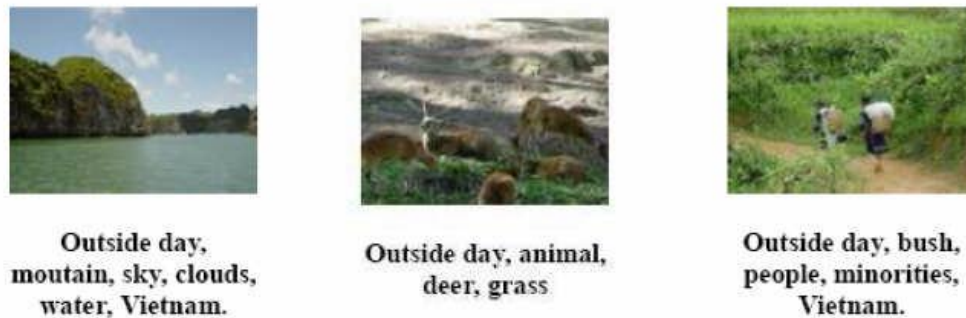


Figure 2.10 Quelques images annotées avec ses mots clés choisis [21]

2.4.3.3 Quelques Systèmes de l'approche orienté-sémantique

Depuis dix ans environ, l'apparition des systèmes de gestion d'images qui ont adopté le paradigme orienté-sémantique a beaucoup augmenté. La plupart de ces systèmes utilisent les mots-clé comme moyen de description d'image. Nous citons :, iBase (<http://www.ibase.com>), Index+ (<http://www.ssl.co.uk>), Digital Catalogue (<http://www.imageres.com>), Fastfoto (<http://www.picdar.co.uk/index2.html>), FotoWare (<http://www.fotoware.com>), etc. [9].

D'autres systèmes plus efficaces existent et utilisent les approches retour de pertinence et description des concepts pour décrire le contenu sémantique des images. IRIS le système de [10], FourEyes du MIT [27]. D'autres systèmes existes aussi tel que :

□ **Le système IM-DBMS** : Dans IM-DBMS, deux bases séparées sont utilisées pour stocker les images : d'une part, la base physique pour le stockage des images, et d'autre part la base logique pour le stockage des descripteurs des images.

□ **Le système ELF** [31] : Est un système relationnel de gestion d'une base de photos d'art. La description de l'image est constituée d'un ensemble d'objets, de leurs caractéristiques et de leurs relations. Basé sur le modèle relationnel, le système ELF utilise les notions :

- **Objets** : permettent de décrire les caractéristiques d'un objet. Par exemple : Maison = (Image, Couleur, Position, etc.).

- **Associations** : permettent de représenter d'une part, la relation de généralisation et, d'autre part, les relations qui existent entre les objets dans une image. Par exemple la relation EST-PRES-DE (FEMME, MAISON, IMAGE1) montre que la femme est près de la maison dans l'image 1.

□ **Le system KMeD (Knowledge-based Multimedia Medical Distributed Database System)** de [18] est un des PACS (Picture Archival and Communication Systems) les plus connus. Il se caractérise par sa capacité à prendre en considération l'évolution du contenu (par exemple une tumeur) de l'image médicale. Les images utilisées dans ce système sont les radiographies et les IRM.

2.4.3.4 Limites de l'approche orientée sémantique

Avec l'intégration de la perception humaine, la description de l'image par le contenu sémantique a donné plus de souplesse à l'expressivité des requêtes. D'ores et déjà, l'utilisateur peut rechercher des images selon sa propre interprétation. Cependant, plusieurs limites se posent :

- L'absence d'une description universelle: Chaque description du contenu sémantique de l'image présente une vue particulière et correspond à un besoin spécifique du domaine d'application. Or, le point de vue de l'expert qui stocke les images n'est pas forcément le même que celui de la personne qui les recherche. La description du sens associé à l'image demeure une tâche assez complexe. Elle dépend de plusieurs facteurs, tels que le domaine d'application, la personne qui regarde l'image, son niveau d'expertise, etc.
- Le coût élevé en temps: La description du contenu sémantique de l'image est souvent manuelle. D'après Eakins [12], l'indexation d'une photographie prend en moyenne 7 minutes.
- L'absence d'une prise en considération de l'évolution du contenu de l'image: Dans certains domaines tels que dans le domaine médical, l'interprétation de certains objets peut évoluer. Les solutions proposées n'ont pas pris en compte cet aspect important dans la description de l'image.

2. 5. Conclusion

Dans cette recherche bibliographique sur la recherche d'images, nous avons exploré, dans un premier lieu, les différents travaux effectués dans ce domaine, et on a traité trois approches qui sont proposées dans la littérature, dont chacun s'évalue en fonction de son niveau d'abstraction par rapport à l'image. Ils sont de trois catégories :

- l'approche orienté-contexte: tente de donner une description plus ou moins précise de l'image par le biais de son contexte.
- l'approche orienté-contenu: apporte une description objective du contenu physique de l'image en manipulant ses couleurs, textures, formes, etc.
- l'approche orienté-sémantique: propose une description de l'interprétation humaine du contenu de l'image.

Pour qu'un système de recherche d'images soit performant et surtout satisfaisant les besoins et les attentes des utilisateurs, l'addition de valeurs sémantique dans le système est indispensable. Traiter les images au niveau sémantique, c'est-à-dire au niveau du jugement que l'humain porte en regardant les images. C'est à cette condition que la recherche d'informations sur les images pourra déboucher sur de réelles avancées répondant à des besoins et applications diverses, tant personnelles que professionnelles.

Dans notre projet nous nous intéressons à travaillé sur l'approche orienté-sémantique c'est à dire sur l'annotation d'image mais de manière automatique.

Pour cela nous parlerons dans le chapitre suivant (chapitre 3) sur l'annotation automatique d'images.

CHAPITRE 3

ANNOTATION AUTOMATIQUE D'IMAGE

3.1 Introduction

L'annotation automatique d'images est un domaine du traitement d'images permettant d'associer automatiquement des mots-clés ou du texte à des images à partir de leur contenu afin de pouvoir ensuite rechercher des images par requête textuelle.

Dans ce chapitre on va faire un survol sur des modèles et des méthodes déjà utilisés dans l'annotation.

Dans ce chapitre, nous rappelons tout d'abord l'annotation d'images et leurs types. Ensuite, nous allons introduire quelques travaux existants sur l'annotation automatique d'images puis nous présentons une architecture générale d'une approche d'annotation automatique d'image et leurs détails, ensuite nous présentons quelques modèles d'annotation automatique d'images qu'il existe dans la littérature.

3.2 Définition de l'annotation d'images

Nous avons trouvé différentes définitions informelles qui dépendent du domaine de recherche, D'abord dans le contexte des Interfaces Homme Machine [38], les auteurs définissent une annotation comme 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.

Selon les psycholinguistes et cognitivistes, l'annotation est une trace de l'état mental du lecteur et une trace de ses réactions vis-à-vis du document [39]. Elle est concrétisée par des

marques reflétant l'intérêt du lecteur, son activité. On parle alors de « lecture active » [40] par opposition à une lecture de loisir [41]. Annoter est donc une tâche cognitive plus intense que la lecture simple.

Le point de vue des documentalistes [42] définit l'annotation comme l'activité du lecteur qui consiste à poser des marques graphiques ou textuelles sur un document papier, et ce suivant plusieurs objectifs. On retrouve dans cette définition l'annotation comme résultat d'une activité du lecteur, qui précise de plus qu'elle répond à des objectifs.

Selon [Bringay, Barry, Charlet][43] : une annotation est une note particulière attachée à une cible. La cible peut être une collection de documents, un document, un segment de document (paragraphe, groupe de mot, mot, image ou partie d'image, etc.), une autre annotation. A une annotation correspond au contenu, matérialisé par une inscription, qui est une trace de la représentation mentale que l'annotateur se fait de la cible. Le contenu de l'annotation pourra être interprété ç son tour par un autre lecteur. Nous appelons l'ancre ce qui lie l'annotation à la cible (un trait, un passage entouré, etc.).

Selon le dictionnaire Larousse, l'annotation est l'action de faire des remarques sur un texte pour l'expliquer ou le commenter.

Finalement, l'annotation d'images est un processus qui permet d'attribuer à une image un mot ou une liste de mots qui décrivent la sémantique de cette image. Généralement utilisée dans les systèmes de recherche d'images pour retrouver des images pertinentes dans une base de données d'image. On trouve dans la littérature de nombreux ouvrages sur l'annotation manuelle, automatique ou semi-automatique d'images.

3.2.1 Annotation manuelle

L'annotation manuelle est un processus qui consiste à attribuer manuellement à chaque image un ensemble de mots clé. Ce processus est généralement réalisé par un expert humain (par exemple : un médecin ophtalmologue annote les images rétiniennes).

Dans ce type d'annotation d'image, la précision augmente (expertise humain) et la productivité baisse (lorsque le volume de la base d'images devient important).

3.2.2 Annotation automatique

L'annotation automatique, ou auto-annotation, s'effectue à travers un système informatique (machine) sans intervention humaine. C'est une indexation textuelle automatique à partir du contenu visuel de l'image.

Dans ce type d'annotation d'image, La précision est faible par rapport à l'annotation manuelle, mais la productivité augmente (possibilité d'annoter un volume important d'images).[44]

3.2.3 Annotation semi-automatique

L'annotation semi-automatique est la combinaison des deux types d'annotations précédentes. Généralement un processus d'annotation automatique est initié, suivi par une validation d'un expert humain.

3.3 Etat de l'art sur l'annotation d'image

Les premières publications sur l'annotation automatique d'images sont apparues récemment, vers 1999, et depuis ce domaine a suscité beaucoup d'intérêt dans la communauté du traitement d'images. Dans sa thèse, S. Tollari [45] (p96) a dressé un tableau comparatif d'une vingtaine de modèles d'annotation automatique proposés entre 1999 et 2006. Nous incluons de nombreux autres travaux dans cette section.

[Maron, Ratan][46] ont abordé l'annotation automatique mais en utilisant un seul mot clé à la fois. Mori et al., [47] 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 al., [48] ont proposé un modèle, appelé "modèle de translation", qui est une amélioration du modèle de co-occurrence proposé par Mori et al., [47], en utilisant un algorithme d'apprentissage. Des modèles probabilistes tels que "Cross Media Relevance model" [49] et "Latent Semantic Analysis" [50] ont été aussi proposés.

Jia et Wang [51] utilisent les chaînes de markov cachées bidimensionnelles pour annoter les images.

R. Albatal, [52] propose une méthode d'annotation d'images basée sur les Phrases Visuelles (Une Phrase Visuelle est un ensemble de régions d'intérêts construit suivant des critères prédéfinis; un critère proposé et étudié ici est de nature topologique.)

H. Bannour, C. Hudelot [53] propose une nouvelle méthode pour la construction automatique de hiérarchies sémantiques adaptées à la classification et à l'annotation d'images, La construction de la hiérarchie est basée sur une nouvelle mesure de similarité sémantique qui intègre plusieurs sources d'informations : visuelle, conceptuelle et contextuelle. L'objectif est de fournir une mesure qui est plus proche de la sémantique des images.

Dans [54], les auteurs [L. Thi] ont construit un système de recherche d'images par le contenu sémantique. Cependant, ce n'est pas un système d'apprentissage. Ce système utilise une base d'images qui a plus de 3500 images indexées manuelles avec des mots-clés.

Dans [55], Les auteurs [G. Carneiro, N. Vasconcelos] ont construit un système qui permet de faire l'annotation d'images en utilisant l'apprentissage supervisé. Dans [56], afin de faire l'annotation d'images les auteurs ont obtenu quelques résultats en utilisant le modèle FACMRM (the Fixed Annotation-based Cross-Media Relevance Model).

Dans [57], l'auteur [Nguyen] a construit un système de recherche d'images sémantique. Pour chaque image, il la segmente en des régions. Puis, il calcule des caractéristiques de chaque région. Pour connecter les concepts et le caractéristique bas niveau de région, il utilise le réseau de neurone et l'entraîne pour classifier des régions dans une image.

S. Barrat, S. Tabbone [58], présente une nouvelle méthode qui améliore l'efficacité de la classification d'images par le contenu, en intégrant des concepts sémantiques extraits du texte, et en étendant automatiquement les annotations existantes à des images non annotées ou faiblement annotées. Ce modèle s'inspire de la théorie des modèles graphiques probabilistes, permettant de traiter les données manquantes.

K. Barnard, Duygulu [59] segmentent les images en régions. Chaque région est représentée par un ensemble de caractéristiques visuelles et un ensemble de mots-clés. Les images sont alors classifiées en modélisant de façon hiérarchique les distributions de leurs mots-clés et caractéristiques visuelles.

Grsky et al. [60] associent des coefficients aux mots afin de réduire la dimensionnalité. Les vecteurs de caractéristiques visuelles et les vecteurs d'indices correspondant aux mots-clés sont concaténés pour procéder à la recherche d'images.

Benitez et al. [61] extraient de la connaissance à partir de collections d'images annotées, en classifiant les images représentées par leurs caractéristiques visuelles et textuelles.

Hakim Hacid [62] présente un travail d'annotation semi-automatique d'images basée sur les graphes de voisinage. Pour l'assignation d'une sémantique aux images. Il considère l'annotation semi-automatique car elle nécessite l'intervention de l'utilisateur pour valider les décisions du système.

3.4 Architecture générale d'une approche d'annotation automatique d'image

Cette approche d'image est basée sur un apprentissage supervisé, elle comprend deux phases principales (voir la figure 3.1)

1. Une *phase d'apprentissage* chargée d'apprendre les relations entre des descripteurs visuels (signaux visuels de bas niveau) et des annotations symboliques associées à ces descripteurs. L'apprentissage est effectué en s'appuyant sur des collections d'images annotées (corpus) sur lesquelles une extraction de contenus visuels est effectuée. Cette phase génère une base des blobs (classes) annotés.
2. Une *phase de reconnaissance* prenant comme entrée des images non annotées (nouvelle image), et propose des annotations symboliques en s'appuyant sur les contenus visuels des images d'entrée et la base des blobs (classes) annotés appris dans la phase d'apprentissage.

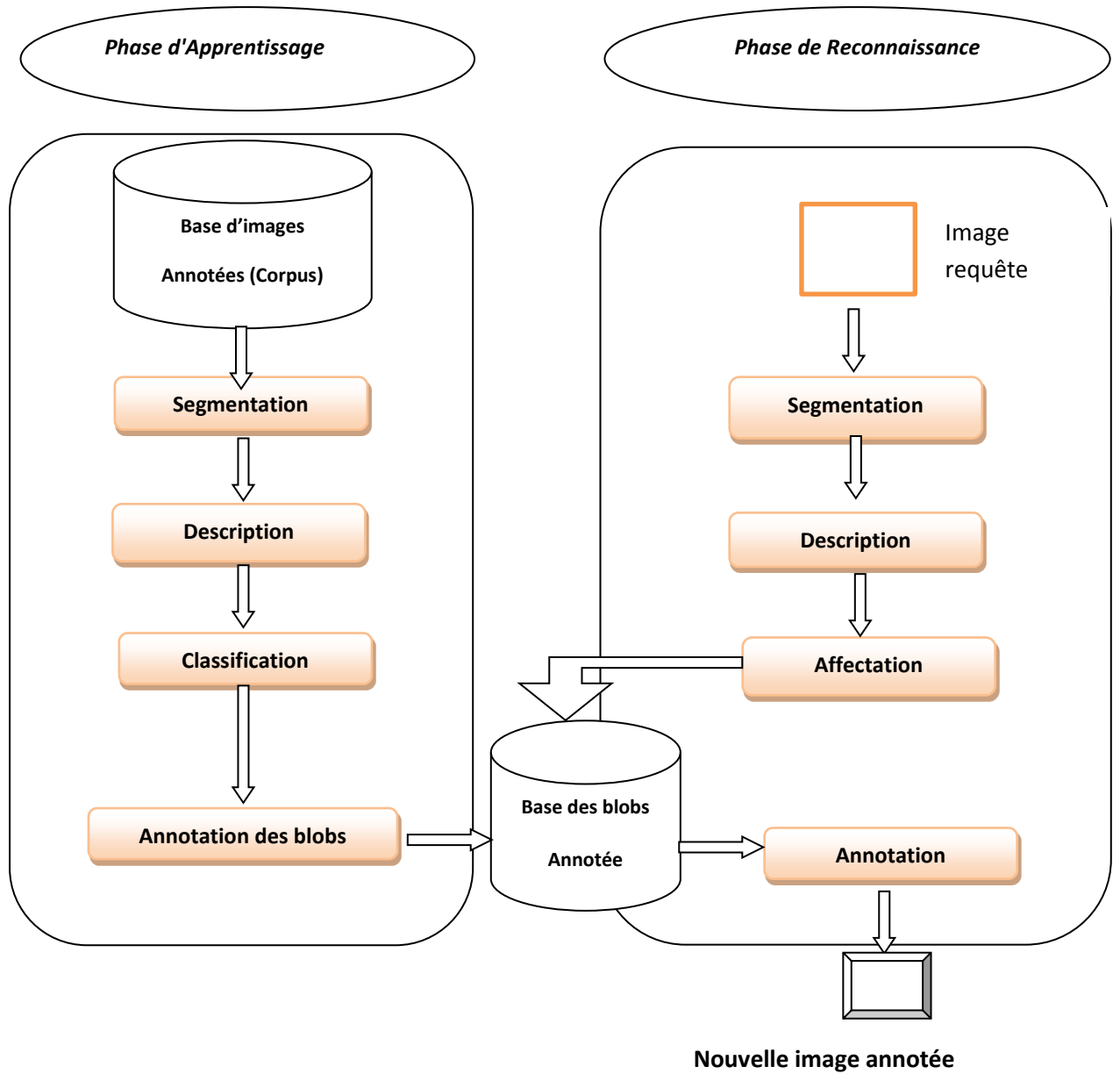


Figure 3.1 Architecture générale d'une approche d'annotation automatique d'image.

3.4.1 Extraction du contenu visuel

L'extraction du contenu visuel est généralement constituée de deux étapes successives :

1. La segmentation en régions : le but de cette étape est de déterminer des parties (ensembles de pixels connexes) de l'image supposées contenir des informations visuelles pertinentes et utiles pour l'annotation.
2. La description des régions : dans cette étape, une description des informations visuelles des régions segmentées est effectuée. Le résultat de la description d'une région est appelé descripteur visuel de région.

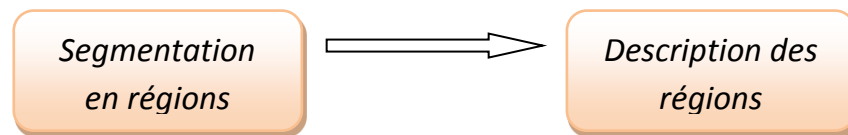


Figure 3.2 Etapes d'extraction du contenu Visuel d'image.

Le processus d'apprentissage supervisé est effectué sur des descripteurs du contenu visuel associés à des symboles d'annotation. La figure 3.3 montre un exemple d'extraction du contenu visuel d'une image. Dans cette figure, une image est d'abord segmentée en plusieurs régions, ensuite chaque région est décrite par un descripteur, et enfin les descripteurs des régions sont synthétisés dans un descripteur de contenu visuel.

Dans la suite, nous présentons un état de l'art sur chacune de ces trois étapes.

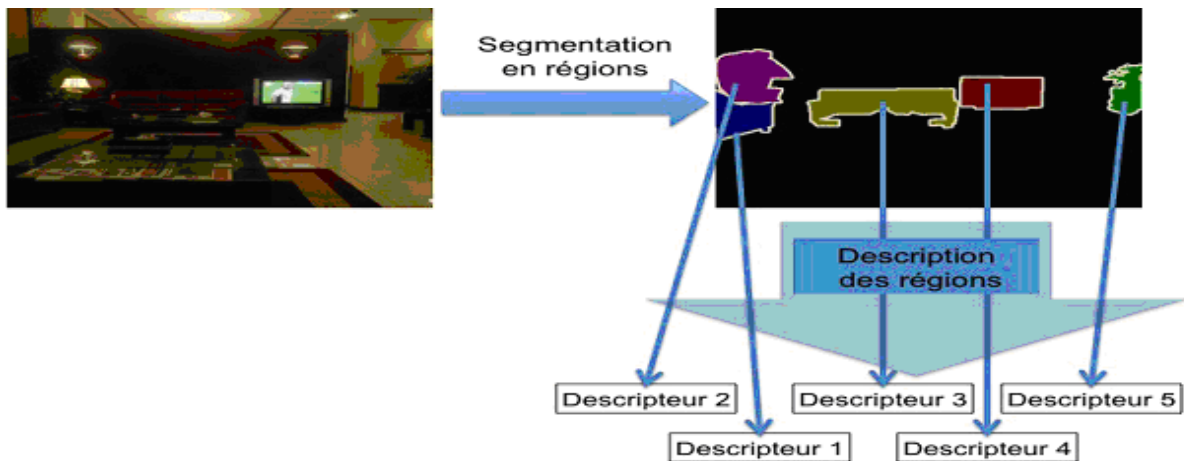


Figure 3.3 Extraction du contenu visuel d'une image.

3.4.1.1 Segmentation en régions

La segmentation d'image est la première étape et aussi l'une des tâches les plus critiques de l'analyse d'image. Il est évident que les résultats de la segmentation ont une influence considérable sur cette analyse et la compréhension de l'image [63].

Dans le domaine d'annotation automatique d'images, l'hypothèse qui sous-tend la segmentation est que des parties spécifiques de l'image aident à la description visuelle des objets à détecter dans les images.

La segmentation est la partition d'une image en un ensemble de régions qui ne se chevauchent pas et dont l'union est l'image entière (voir figure 3.4). Quelques règles à suivre pour obtenir une segmentation sont :

1. Les régions doivent être uniformes et homogènes par rapport à certaines caractéristiques (niveau de gris, écart type, gradient).
2. Leurs intérieurs doivent être simple et sans beaucoup de petits trous (des parties de région non segmentés).
3. Les régions adjacentes doivent avoir des valeurs très différentes par rapport à la caractéristique prise en compte dans la segmentation.
4. Les limites de chaque région doivent être simples et spatialement précises

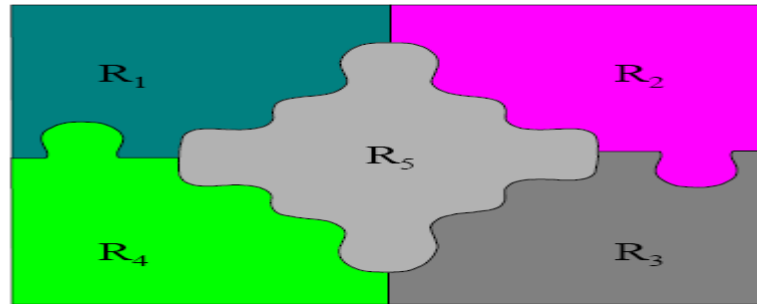


Figure 3.4 Exemple de segmentation d'image.

Plusieurs techniques de segmentation existent dans la littérature, comme : le découpage régulier, la segmentation adaptative et la segmentation en régions d'intérêt.

- **Différentes approches de segmentation**

Beaucoup de méthodes de segmentation sont basées sur deux propriétés de base des pixels par rapport à leur voisinage local : *discontinuité* et *similitude*. La discontinuité de pixels est utilisée par les méthodes de segmentation par approche contours, tandis que la similitude de pixels est utilisée par les méthodes de segmentation par approche région.

La figure ci-dessous (figure 3.5) montre les différentes approches de segmentation.

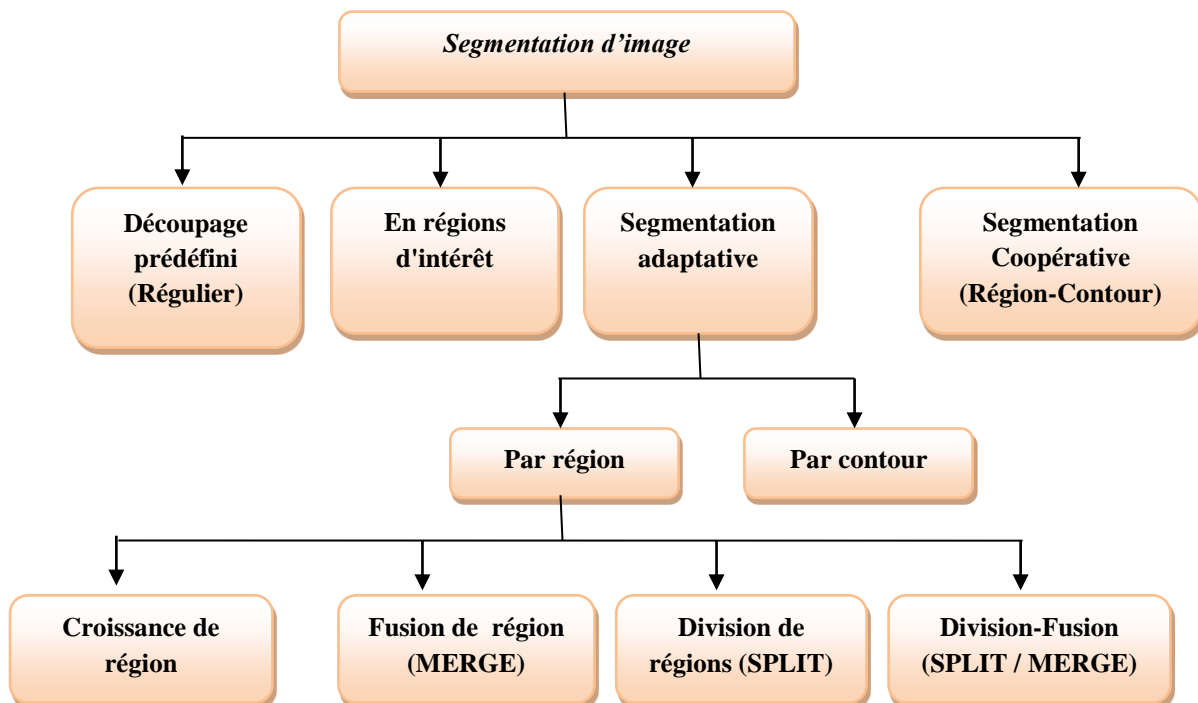


Figure 3.5 Les approches de segmentation d'image.

3.4.2 Classification

Dans le but de réaliser un système de regroupement (clustering) d'images en familles d'images visuellement similaires, cette étape permet de regrouper des régions visuellement similaires en classes appelées (blobs).

Parmi toutes les techniques de classification, on distingue :

- La classification *supervisée* dans laquelle un expert a fourni le modèle exact des classes à obtenir.
- La classification *non supervisée* dans laquelle le nombre de classes, inconnu a priori, est déduit directement des données.

On distingue ensuite :

- La classification *avec apprentissage* dans laquelle on entraîne le classifieur à l'aide d'un ensemble de données connues a priori. Cet entraînement a pour but d'adapter les sorties du classifieur en fonction des entrées qu'on lui soumet.
- La classification *sans apprentissage* où le classifieur travaille directement sur les données sans aucune connaissance préalable.

On trouve enfin des méthodes de classification :

- *paramétriques* dans lesquelles on fournit un certain nombre de paramètres au classifieur qui vont influencer le résultat de la classification.
- *non-paramétriques* dans lesquelles le classifieur doit se débrouiller seul pour classer les données sans aide extérieure.

Dans le cas de la classification supervisée, un expert identifie les classes de données et on classe les données existantes parmi elles. Dans le cas non supervisé, les classes sont construites en fonction des données, selon l'algorithme de classification employé.

Parmi les méthodes de classification par apprentissage supervisé les plus utilisées dans la littérature pour le problème de la classification d'images : les k- plus proches voisins

(K-means), la classification bayésienne, les machines à vecteurs de support (SVM), les réseaux de neurones, et les arbres de décision.

3.4.3 Annotation des blobs

Cette étape consiste à corrélérer le contenu visuel (bas niveau) avec le contenu sémantique (haut niveau) des régions des images en annotant les blobs regroupant les régions visuellement similaires, par les mots clés d'images de la base de données traitée.

Pour cela, il existe plusieurs méthodes d'annotation d'images sous forme de modèles existantes dans la littérature (pour plus de détail, voir la section suivante).

3.5 Modèles d'annotation automatique d'images à partir du contenu visuel

Il existe de très Grand nombre de modèles :

- Modèles basés sur l'analyse de la sémantique latente (LSA, PLSA...)
- Modèles basés sur la distribution de Dirichlet (MoM-LDA, Corr-LDA...)
- Modèles de traduction de langues...

3.5.1 Modèles basés sur l'analyse de la sémantique latente

Cette technique utilise tout d'abord pour l'analyse de grand corpus de textes, LSA [64] pour Latent Semantic Analysis ou analyse de la sémantique latente, est une technique statistique automatique pour extraire et inférer des relations entre mots à partir de leurs contextes.

L. Cascia, M. Sethi [65], un premier système de recherche d'image sur le web utilisant le modèle LSA appelé ImageRover est proposé, a pour objectif de combiner les informations textuelles d'une page web avec le contenu visuel des images. Pour les attributs visuels des histogrammes de couleurs et de textures sont extraits, ImageRover intègre aussi des caractéristiques textuelles pour mieux exprimer le contenu sémantique de l'image. Pour chaque image d'une page web (document HTML) un vecteur de mots est créé, en fonction de

leurs fréquences d'occurrences dans la page web. Le système assigne des poids différents aux mots par exemple les mots dans le titre ou dans les champs ALT et IMG sont pondérés plus fortement, en plus les poids associés aux mots qui apparaissent avant ou après une image sont calculés en fonction de leurs positions par rapport à l'image et aussi leurs distance de l'image.

Chaque image est finalement indexée par un vecteur global concaténant les vecteurs visuels et textuels. Une matrice terme-image est construite, puis décomposée par SVD. La recherche d'image revient alors à une recherche par plus proche voisin. L'utilisateur peut réaliser une requête par mots-clés, transformée par le système en un document-requête, la recherche par plus proche voisin est effectuée dans l'espace textuel latent, il peut aussi choisir plusieurs images-requêtes. Les deux types de requêtes pouvant être combinées et raffinées par le bouclage de pertinence.

Des expériences montrent que la méthode par combinaison de texte et d'image avec bouclage de pertinence est plus efficace que la méthode de recherche par mots-clés seuls, ou par mots-clés et bouclage de pertinence, et par contenu visuel et bouclage de pertinence, cependant, comme LSI est utilisé seulement sur le texte, le modèle n'est pas capable de trouver les cooccurrences entre les traits visuels et les mots.

Dans [66], LSA et PLSA sont utilisées pour construire un système d'auto annotation. dans leur modèles LSA et PLSA-MIXED, chaque image de la base d'apprentissage est représentée par un vecteur concaténant un vecteur textuel de 149 dimensions et un vecteur de couleur RVB de 648 dimensions. Les expériences qu'ils ont menées montre que dans leurs cas, le modèle LSA est meilleur que le modèle PLSAMIXED. Ils supposent dans [70] que la raison du mauvais score de PLSA MIXED et que les deux modalités (textuelle et visuelle) sont définies dans PLSA MIXED avec la même importance lors de la définition de l'espace latent. Or la modalité visuelle contient moins d'information sémantique. c'est pourquoi ils proposent dans [67] de construire un espace latent pour chacune des modalités, ce modèle PLSA-WORDS est construit ainsi :

1. Les probabilités $p(w/z)$ et $p(z/d)$ sont apprises uniquement sur les mots clés des images
2. Un deuxième modèle PLSA est appris sur des traits visuels pour $p(v/z)$ mais en gardant la probabilité $p(z/d)$ apprise sur les mots clés ;
3. Soit une nouvelle image d_{new} que l'on veut annoter (on connaît seulement les paramètres visuels), on peut estimer $p(z/d)$ à l'aide de la probabilité $p(v/z)$ calculé en (2) et de l'algorithme EM (Estimation- Maximisation)
4. la probabilité a posteriori des mots clés pour l'image d_{new} est inférée par :

$$P(W/d_{new}) = \sum_{z_k \in Z} P(W/Z_k) * P(Z_k/d_{new}) \quad (3.1)$$

3.5.2 Modèles fondés sur la distribution de Dirichlet

La distribution de Dirichlet estime le vecteur de probabilité $\theta^d = \{p_1, p_2, \dots, p_r\}$ où p_j est la probabilité que le concept (appelé aussi classe cachée ou latente) z_j soit dans le document d , en fonction du nombre d'occurrences α_j de chaque concept dans le document.

Il existe plusieurs modèles de Dirichlet, parmi eux nous citons : LDA, GM-Mixture, GM-LDA, MOM-LDA, Corr-LDA. Ces modèles sont détaillés dans [68].

3.5.3 Modèle d'Annotation d'image par la fouille de donnée

La combinaison des données textuelles (symbole ou concept) avec celles qui proviennent du signal (caractéristiques visuelles) dans un but d'annotation automatique des images est discutée dans [69]. La technique des règles d'associations du processus de fouille d'image permet de répondre précisément à ce besoin.

Les résultats de ces expérimentations montrent qu'un processus de fouille (FDI) peut être utilisé pour combiner deux sources de données différentes, ils sont jugés que la FDI est alors une perspective intéressante pour l'indexation automatique des images. Ces résultats montrent la faisabilité et l'efficacité de cette approche.

Wang, Zhang, Jing [70], il a proposé un système nommé AnnotSearch qui annote les images en utilisant les techniques de datamining. Cette approche est composée de trois étapes : la première étape permet de faire la recherche basée sur le texte par une requête textuelle, la

deuxième étape permet de faire la recherche basée sur le contenu visuel par une requête image et la troisième étape permet de catégoriser les images trouvées et à extraire les formules marquantes depuis leurs descriptions.

Les résultats de ces expérimentations montrent l'efficacité et le bon fonctionnement de ce système.

3.6 Conclusion

Dans ce chapitre, nous avons présenté l'annotation d'images et leurs types qui sont : l'annotation manuelle, automatique et semi-automatique. Ensuite, nous avons présenté quelques travaux existants sur l'annotation automatique d'images puis nous avons présentés une architecture générale d'une approche d'annotation automatique d'image et leurs détails, ensuite nous avons présenté quelques modèles d'annotation d'images.

Dans le chapitre qui suit, nous aborderons notre contribution qui vise à améliorer l'annotation automatique d'images en utilisant l'arbre quaternaire et la partie de parallélisation dans ce processus pour gagner du temps.

CHAPITRE 4

APPROCHE PARALLELE D'ANNOTATION AUTOMATIQUE D'IMAGES

4.1 Introduction

Ce chapitre décrit l'approche adoptée pour atteindre les objectifs fixés, en tenant compte des spécificités des lésions caractéristiques de la pathologie étudiée, en effet, dans plusieurs domaines d'application de l'annotation d'images.

Dans le domaine médical, des descripteurs locaux stockant les caractéristiques visuelles de régions d'image, sont nécessaires car le nombre de pixels représentant une pathologie est faible par rapport au nombre total de pixels dans l'image.

Du fait que notre objectif est d'annoter une image médicale dans le but de détecter des pathologies, cela nécessite une indexation locale à un niveau de détail augmenté, pour cela la structure d'arbre quaternaire à trois niveaux répond à cet objectif.

Dans ce chapitre, nous exploitons les descripteurs visuels de la norme MPEG7 (Moving Picture Experts Group version 7) de couleur, texture et de forme. Puis nous introduisons la notion de parallélisme dans l'approche d'annotation en utilisant le réseau informatique.

4.2 Concept de base de l'arbre quaternaire

4.2.1 Définition

L'arbre quaternaire (QuadTree) est très utilisé dans le domaine des images, aussi bien pour le stockage, la compression, l'extraction de l'information sur le contenu des images, que pour la recherche d'image par le contenu. L'arbre quaternaire est une structure hiérarchique construite par divisions récursives de l'espace en quatre quadrants disjoints [71]. Cette forme de structure est très utilisée pour modéliser les images [72] ou pour indexer les

caractéristiques des images. En particulier, l'arbre permet de tenir compte, lors de la phase de recherche d'images similaires, de la localisation spatiale des caractéristiques d'images [73]. Pour cela, une image est récursivement décomposée en quatre quadrants disjoints de même taille, en fonction d'un critère de découpage de telle sorte que chaque nœud de l'arbre représente un quadrant dans l'image. Le nœud racine représente l'image entière.

Lors de l'étape de la construction de l'arbre quaternaire, chaque nœud de l'arbre est identifié par un numéro séquentiel correspondant au identifiant de nœud père et en suivant une direction Nord-Est, Nord-Ouest, Sud-Est, Sud-Ouest.

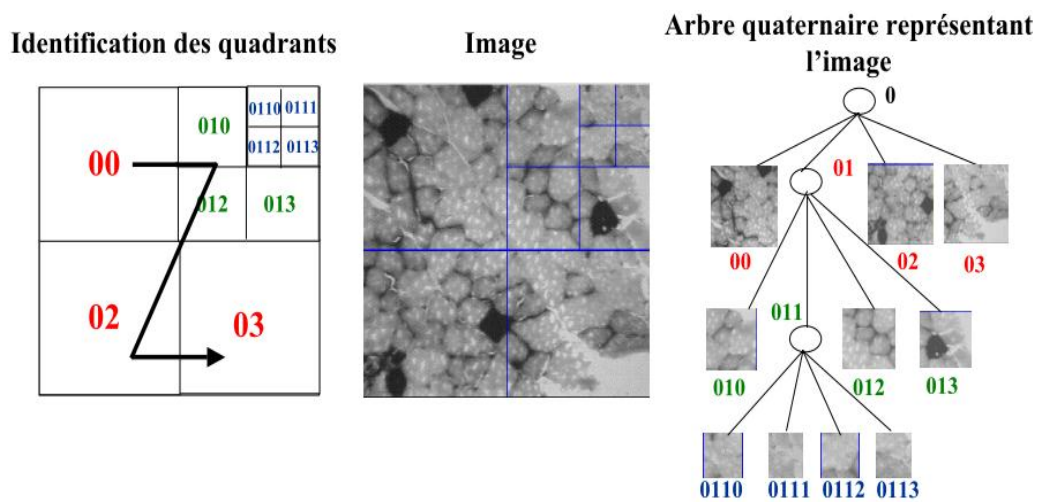


Figure. 4.1 Représentation d'une image par l'arbre quaternaire.

Il existe plusieurs fonctions permettant d'associer un identificateur à un nœud d'arbre quaternaire [71]. Ces fonctions permettent de retrouver facilement, à partir de l'identificateur de l'image et du nœud d'arbre quaternaire, le quadrant associé dans l'image.

Si on suit un ordre en Z, en suivant les directions Nord-Ouest, Nord-Est, Sud-Ouest et Sud-Est et en leur associant respectivement les identificateurs 0,1,2,3. Pour commencer l'entier 0 identifie l'image entière et le nœud racine de l'arbre quaternaire, les entiers 0 à 3 précédé de 0, identifient les quatre premiers quadrants de l'image et les quatre nœuds quaternaire associés. De manière récursive, les nœuds fils d'un nœud identifié par k sont identifiés par kx avec x un entier prenant sa valeur dans $[0,3]$. Deux nœuds de même identificateur dans deux arbres quaternaires différents sont nommés *nœuds homologues* [73].

4.3 Similarité d'images représentées par des arbres quaternaires [73]

4.3.1 Distances entre nœuds d'arbre quaternaires

A chaque nœud d'arbre quaternaire correspond un quadrant dans l'image associée. Un nœud d'arbre quaternaire peut contenir n'importe quelle information sur la région correspondante : un histogramme de couleurs, un vecteur de caractéristiques des points d'intérêt, la sous-image (compressée ou non) correspondant à la région, ou des descripteurs de Fourier, par exemple. Il est par conséquent possible de calculer des distances entre nœuds d'arbre quaternaires.

On note $\delta_k(i, j)$ est une distance normalisée ($\delta_k(i, j) \in [0, 1]$) entre deux nœuds homologues k apparaissant dans les arbres quaternaire i, j la distance δ_k peut-être n'importe quelle distance applique aux nœuds d'arbre quaternaire en fonction du contenu des nœuds [74].

4.3.2 Définition générale de la distance Δ entre images

La distance Δ est une distance entre images représentées par des arbres quaternaires. Sa définition est générale et permet de distinguer plusieurs familles de distances entre arbres quaternaires, la distance Δ entre deux images i et j est définie par une somme de distances $\delta_k(i, j)$ entre les nœuds des arbres quaternaires représentant les images i et j pondérées par des coefficients $c_k, c_k > 0$:

$$\Delta(i, j) = \frac{\sum_k c_k \delta_k(i, j)}{\sum c_k} \quad (4.1)$$

- ✓ k est l'identificateur d'un nœud pris parmi l'union des identificateurs de nœuds apparaissant dans les arbres quaternaires des images i et j , On note N a cardinalité de l'ensemble des identificateurs de nœuds.
- ✓ $\delta_k(i, j)$ est une distance normalisée entre les nœuds homologues k des arbres quaternaires i, j .
- ✓ c_k est un coefficient positif représentant le poids du nœud k dans le calcul de la distance Δ . Le choix de chaque poids c_k dépend des besoins de l'utilisateur, c'est-à-dire de l'importance que l'utilisateur souhaite donner à certains quadrants d'image par rapport à d'autres dans le calcul de la distance, où $c_k = 4^{-p}$, p est la profondeur de nœud dans l'arbre quaternaire.

4.3.3 Descripteurs utilisés

Le contenu de l'image désigne à la fois son contenu sémantique et son contenu visuel. Le contenu visuel est décrit par la texture, la forme et la couleur [75]. Un descripteur est défini comme la connaissance utilisée pour caractériser l'information contenue dans les images. De nombreux descripteurs sont utilisés dans les systèmes de recherche pour décrire les images de détail des descripteurs de bas niveaux utilisés dans la littérature est détaillés dans [75].

Ces attributs (couleur, forme, texture) vont représenter l'image, leurs choix est déterminant pour la qualité des résultats. Si les attributs sont mal choisis, la méthode d'Indexation/Recherche donnera de mauvais résultat.

Ces attributs sont stockés dans des vecteurs descripteurs ; Il est important de savoir que l'utilisation d'un seul critère (couleur, Forme ou Texture par exemple) ne suffit généralement pas pour obtenir de bons résultats. La combinaison de plusieurs critères semble plus appropriée.

Dans notre approche, nous avons retenu la couleur, la texture et la forme pour caractériser l'image de la rétine et la distance est une distance multi-niveaux. Les descripteurs utilisés sont définies comme suit :

a. Détection des contours de l'image

Les contours constituent des indices riches pour toute interprétation ultérieure de l'image. Dans notre approche on utilisant la détection de contours par filtrage basée sur le filtre de sobel. Cette technique utilise l'approche gradient qui fait intervenir la convolution d'image à partir de filtres séparables. La convolution d'image est considérée l'image comme une matrice dont chaque coefficient représente un point de l'image. Chaque point de l'image a donc un indice de ligne ainsi qu'un indice de colonne.

Pour obtenir les contours dans l'image, il suffit donc d'appliquer sur celle-ci deux filtres matriciels (selon les axes horizontaux et verticaux) et de calculer une norme.[76]

L'opérateur de Sobel permet d'estimer localement la norme du gradient spatial bidimensionnel d'une image en niveau de gris. Il amplifie les régions de fortes variations locales d'intensité correspondant aux contours. La norme du gradient ainsi estimée correspond à l'intensité attribuée au pixel courant. C'est donc l'image de la norme du gradient que l'on visualise généralement.

b. Histogramme d'orientation des contours (EHD)

Pour extraire le vecteur descripteur qui caractérise l'image rétinienne basée sur la texture, nous utilisons l'histogramme d'orientation des contours (Edge Histogram Descriptor) qui est retenu dans la norme MPEG7.

Il représente la distribution spatiale de 5 types de contours pour 16 régions locales d'une image donnée. Il s'agit de quatre contours directionnels (horizontal (0°), vertical (90°) et deux diagonaux (45° et 135°)) et un contour non directionnel (orientation non spécifique) pour chaque région de l'image. La distribution de contours ne constitue pas seulement une bonne représentation de texture mais elle est aussi utile pour la comparaison des images en absence de toute texture homogène.[77]

c. Scalable Histogramme de couleur (SCD)

Le fort pouvoir de discrimination de la couleur en fait un attribut très important dans la grande majorité des systèmes d'indexation et de recherche par le contenu. De nombreux descripteurs sont proposés dans la littérature. L'approche la plus courante et la plus rencontrée dans la littérature est l'histogramme couleur.

Dans notre travail on utilisant Le descripteur correspondant à l'histogramme scalable de couleur (SCD) de la norme MPEG7.

4.4 Approche proposée en annotation automatique d'image basée sur l'arbre quaternaire

Etape1 : dans cette étape, Pour chaque image de la base de données déjà annotées, nous construisons leur arbre quaternaire équilibré à trois niveaux. A chaque nœud de l'arbre quaternaire correspond un quadrant dans l'image associé. Un nœud d'arbre quaternaire peut contenir n'importe quelle information sur la région correspondante : un histogramme de couleurs, un vecteur de caractéristiques, des points d'intérêt, ou des descripteurs de Fourier, par exemple.[78] nous avons stocké les informations suivantes pour chaque quadrant d'un niveau 'n' tel que $n=0, 1, 2$.

- Le niveau 0 : le descripteur de forme (VD(F)) : filtre de sobel, et les mots clés des pathologies rétiniennes qui annotent l'image (N : Normal, P: Micro-anévrisme)

- Le niveau 1 : le descripteur de texture (VD(T)) : EHD
- Le niveau 2 : le descripteur de couleur (VD(C)) : SCD

Le schéma suivant (voir la figure 4.2) montre une telle structure proposée.

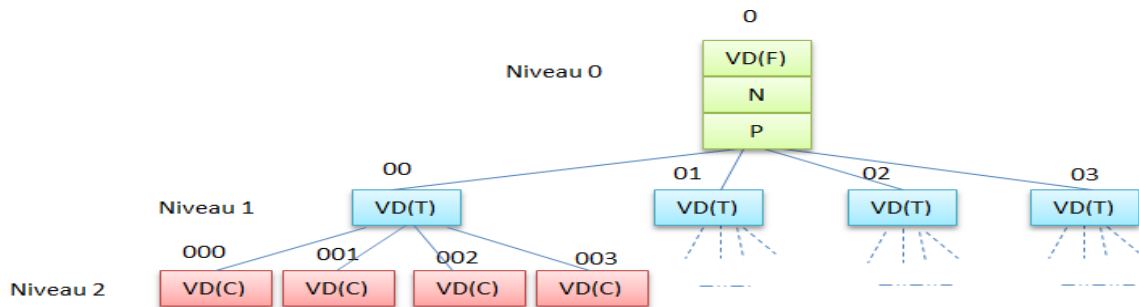


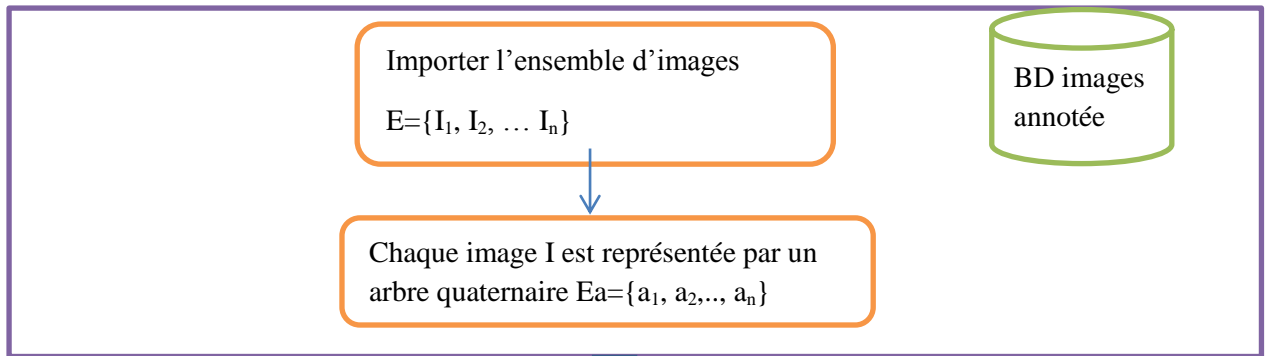
Figure 4.2 La structure d'arbre quaternaire équilibré proposée pour notre système d'annotation automatique.

Etape2 : Algorithme d'annotation proposé

Dans cette étape, nous allons proposer un algorithme d'annotation automatique exploite la recherche d'image par le contenu basée sur la méthode d'arbre quaternaire équilibré à trois niveaux.

Cette approche est divisée en 2 phases principales : phase offline (apprentissage) et phase online (recherche et annotation). La figure 4.3 schématise les phases du système d'annotation automatique d'image proposé.

Phase offline (Apprentissage)



Phase online (Recherche)

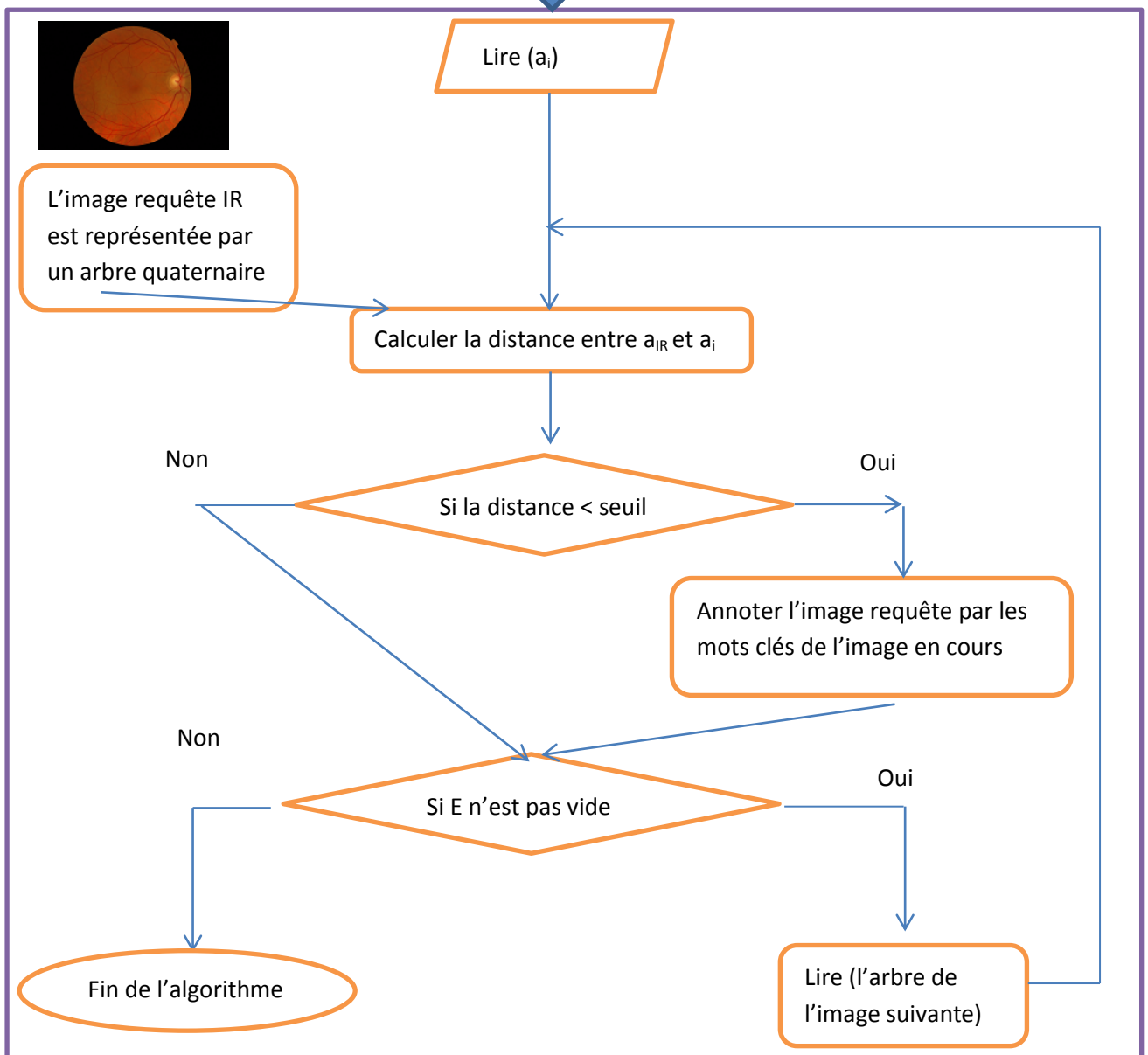


Figure 4.3 Annotation Automatique d'image par arbre quaternaire

4.4.1 Phase offline

Construction de l'arbre quaternaire à 3 niveaux pour toutes les images de la base d'images rétinienne déjà annotées.

Niveau 0 : L'image entière (vecteur descripteur) se trouvant dans ce niveau (c'est la racine de l'arbre quaternaire) ainsi les mots clés indiquant la pathologie de l'image (micro-anévrismes ou normale)

Niveau 1 et 2 : pour chaque nœud, nous stockons les vecteurs descripteurs de chaque quadrant après décomposition en 4 quadrants.

4.4.2 Phase online (Recherche et annotation)

- L'image requête (image non annotée) est représentée par l'arbre quaternaire équilibré à 3 niveaux.
- Comparaison entre l'arbre quaternaire de l'image requête et tous les arbres quaternaires de la base en calculant la distance entre arbre (dans notre cas : distance de manevier utilisée [72]).
- Si la distance est inférieure à un seuil choisi, les 2 arbres sont similaires et l'annotation de l'image requête se fait par récupération des mots clés stockés dans le niveau 0 de l'arbre en cours.

4.4.3 Algorithme d'annotation d'image par arbre quaternaire

DEBUT

- Entrée : $E = \{\text{ensemble d'images de la base représenté dont chaque image est représentée par un arbre quaternaire équilibré à 3 niveaux}\}$,

$IR = \text{image requête représentée par l'arbre quaternaire équilibré à 4 niveaux.}$

- Sortie : image requête annotée.

Lire(arbre de l'image à partir de E)

Tant Que E n'est pas vide

faire

Niveau 0 : L'image entière se trouvant dans ce niveau (c'est la racine de l'arbre quaternaire).

D_0 : notée $\delta_k(i, j)$ distance euclidienne entre le vecteur descripteur de la racine de l'arbre de l'image requête et le vecteur descripteur de l'image en cours d'utilisation. Cette distance est pondérée par un coefficient c_k ($c_k > 0$) : dans ce cas $c_k = 1$ (pour le niveau 0, $p=0$);

Niveau 1 : L'image entière est divisée en quatre quadrants (0,1,2,3) suit la direction Z correspond aux quatre nœuds dans l'arbre, nous faisons un calcul de distance $\delta_k(i, j)$ pour chaque quadrant, nous faisons la somme de ces distances puis nous les multiplions par le coefficient (poids) c_k , dans ce cas $c_k = 1/4$ (pour le niveau 1 $p=1$).

Niveau 2 : chaque quadrant du niveau 1 est divisé en quatre quadrants ça donne : $4 \times 4 = 16$ quadrants (donc 16 nœuds au total). Nous faisons la même procédure déjà citée aux niveau 1 pour calculer les distances des 16 quadrants mais avec un coefficient $c_k = 1/16$ (pour le niveau 2 $p=2$).

Pour chaque niveau, nous aurons une distance, et pour obtenir la distance globale entre l'image requête et l'image en cours, nous faisons la moyenne de toutes ces distances obtenues par la formule suivante :

$$\Delta(i, j) = \frac{\sum_k c_k \delta_k(i, j)}{\sum_k c_k} \quad \text{tel que: } i: \text{image en cours, } j: \text{image requête}$$

Pour vérifier la similarité entre l'image requête et l'image en cours, nous faisons un test:

Si la distance Δ est inférieure à un seuil (fixé de manière empirique à partir de plusieurs tests sur la base)

Alors Affecter les mots clés de l'image en cours à l'image requête.

Finsi ;

Lire (l'arbre de l'image suivante) ;

Fait ;

FIN.

4.4.4 Complexité algorithmique

La complexité d'un algorithme permet de calculer la mesure de l'efficacité d'un algorithme pour un type de ressource. Un algorithme à un coût qui est lié au nombre d'opérations effectuées et à l'espace mémoire occupé par les données. Cela revient à calculer la complexité d'un algorithme en temps d'exécution ce qu'on appelle complexité temporelle ou bien en espace mémoire. Dans notre cas nous nous sommes intéressés à la complexité en temps d'exécution. Nous décomposons notre algorithme en quatre étapes, ce qui nous permet d'associer pour chaque étape une complexité algorithmique O_i , tel que : $i= 1$ à 4

4.4.4.1 Calcul de la complexité

Etape 1 (offline) : permet de représenter chaque image rétinienne de la base par un arbre quaternaire à trois niveau, dans chaque niveau on sauvegarde le descripteur associe ainsi que le mot clé indiquant la pathologie. Les descripteurs sont implémentés en boucle, la complexité algorithmique de cet étape est notée par $O_1(N)$

Pour chaque image de la base faire :

- Découper l'image en arbre quaternaire à trois niveaux, le nombre des quadrants vaut 21 ce qui correspond à une complexité : $O_{11}(N)$
- Calculer le descripteur correspond à chaque nœud : $21 * O_{12}(N)$
- Sauvegarder le descripteur dans le quadrant correspond : $O_{13}(1)$
- Sauvegarder le mot clé de la pathologie : $O_{14}(1)$

$$O_1(N) = N \cdot \max(O_{11}(N), 21 * O_{12}(N) + O_{13}(1) + O_{14}(1)) = O_1(N)$$

N est le nombre d'images de la base.

Etape 2 (online) : pour l'image requête nous avons toujours un arbre quaternaire a trois niveau, les descripteurs (contours, texture, couleur) qui ont été implémentés en boucle, on note la complexité de cet étape par : $O_2(N)$.

Le calcul de cette complexité est calculée de la même manière que $O_1(N)$ sauf qu'il y a une seule image au lieu de N images, alors :

$$O_2(N) = \max(O_{11}(N), 21 * O_{12}(N) + O_{13}(N) + O_{14}(N)) = O_2(N)$$

Etape 3 : calcul de la distance, notée par $O_3(1)$. Pour les trois niveaux nous avons la complexité suivante :

Niveau 0 : $1.T_0(N)$

Niveau 1 : $4.T_1(N)$

Niveau 2 : $16.T_2(N)$

Pour le calcul de Q similarité : $T_3(N) = 1.T_0(N) + 4.T_1(N) + 16.T_2(N)$

Ceci se répète (N) fois pour chaque image, donc la complexité de cette étape est :

$$(N)T_3(N) = O_3(N) \quad (N : \text{est le nombre d'images de la base})$$

Etape 4: comparaison de la distance avec le seuil et affectation du mot clé à l'image requête :

$O_4(1)$ Car il y a un traitement indépendant du nombre d'image de la base

La complexité finale de notre algorithme $O(N)$ est égale à :

$$\text{Max}(O_1(N), O_2(N), O_3(N), O_4(1))$$

La complexité de notre algorithme est linéaire $O(N)$. Ceci nous permet d'éviter la complexité exponentielle, qui caractérise les algorithmes intrinsèquement complexes. Dans notre cas si nous avons un processeur pouvant effectuer 10^9 opérations par seconde, nous obtenons les résultats suivants :

- ✓ pour $N = 10^1$ alors $O(N)$ tend vers 0.
- ✓ pour $N = 10^3$ alors $O(N)$ tend vers 0.
- ✓ pour $N = 10^6$ alors $O(N)$ tend vers 0.
- ✓ pour $N = 10^9$ alors $O(N)$ tend vers 1 seconde.
- ✓ pour $N = 10^{12}$ alors $O(N)$ tend vers ∞ .

4.5 Parallélisme d'annotation automatique d'image proposé

4.5.1 Motivation

L'annotation d'image par l'arbre quaternaire (AQ) décompose le problème en sous parties en s'appuyant sur une décomposition en bloc de l'image, cette décomposition découpe l'image en un ensemble d'imagettes qui sont traitées indépendamment. Le problème de la représentation de l'image en AQ se ramène alors en échelle d'un bloc. D'un point de vue technique ce traitement a l'avantage d'être facilement parallélisable. Ce problème consiste à traiter de façon matérielle pour accélérer les temps de traitement. A cette fin, les algorithmes de calcul de distance entre chaque 2 nœuds homologues de 2 arbres quaternaires et même pour la comparaison entre l'arbre de l'image requête et tous les arbres quaternaires des images de la base d'apprentissage doivent être parallélisées pour répartir la charge de travail sur un nombre plus importants d'unités de calcul.

Notre approche consiste à introduire la méthode de parallélisation dans le processus de recherche de requête (comparaison entre l'image requête et chaque image de la base et plus précisément, l'opération de comparaison de l'arbre quaternaire de l'image requête avec tous les arbres quaternaires représentent les images de la base déjà annotée par le calcul de distance δ la (distance Euclidienne) entre chaque 2 nœuds homologues).

Chaque nœud contient le vecteur descripteur suivant notre approche : dans le niveau 0, le nœud racine contient le vecteur descripteur du contour, dans le niveau 1, les nœuds contiennent chacun le vecteur descripteur du texture et dans le niveau 2, les nœuds contiennent chacun le vecteur descripteur du couleur.

La comparaison entre descripteurs est calculé parallèlement entre chaque 2 nœuds homologues, tous ces calculs sont indépendants les uns aux autres. L'arbre quaternaire contient 21 nœuds (21 nœuds pour l'arbre de l'image requête et autant pour l'arbre de l'image de référence).

Chaque arbre contient 3 niveaux :

Niveau 0 : 1 seul nœud.

Niveau 1: 4 nœuds.

Niveau 2 : 16 nœuds.

Donc finalement, nous avons 21 procédures peuvent être exécutées indépendamment et simultanément les uns aux autres, Cela signifie que ces processus sont parallélisés.

Même pour la comparaison entre l'arbre requête avec tous les arbres de la base ce traitement est indépendant et peut être parallélisé.

4.5.2 Architectures parallèles

Dans cette section de ce chapitre, nous allons étudier les différentes architectures parallèles qui existent dans la littérature.

Une machine parallèle consiste en un ensemble de processeurs capables de travailler en coopération afin de résoudre un problème de grosse taille. Elles offrent un potentiel de ressource de calcul important au travers de leurs processeurs, de leurs mémoires, ainsi que de leur bande passante. Cette définition inclut aussi bien les super calculateurs parallèles comportant des centaines ou des milliers de processeurs, que les réseaux de stations de travail ou encore les machines multiprocesseurs.

Les architectures parallèles peuvent être séparées en plusieurs catégories en fonction de la multiplicité des flots d'instructions et de données. Ainsi, Flynn [79] distingue deux types de machines parallèles : *les machines SIMD* (Single Instruction Multiple Data) et *les machines MIMD* (Multiple Instruction Multiple Data). Dans une machine SIMD, tous les processeurs exécutent la même flot d'instructions sur des données différentes. Cette approche permet de réduire à la fois la complexité du matériel et celle du logiciel mais elle est appropriée seulement pour des problèmes caractérisés par un fort degré de régularité.

Dans une machine parallèle MIMD, chaque processeur peut exécuter un flot d'instructions sépare sur ses propres données locales.

Les machines SIMD sont plus simples à utiliser mais elles ne permettent cependant pas de traiter efficacement tous les types de problèmes. Les machines MIMD ont le mérite d'être d'un usage plus général et d'avoir un cout moins élevé. C'est pourquoi la plupart des ordinateurs parallèles sont aujourd'hui des machines MIMD.

Les machines MIMD sont elles-mêmes subdivisées en deux groupes [80], les multi ordinateurs et les multiprocesseurs, qui se différencient par la localisation de leur mémoire.

Les multiprocesseurs, également appelés *machines MIMD à mémoire partagée*, sont caractérisés par plusieurs processeurs partageant le même espace d'adressage. Tandis que dans le groupe des multi ordinateurs ou *machines à mémoire distribuée*, les nœuds qui définissent la machine parallèle sont indépendants les uns des autres, chaque nœud étant composé d'un processeur et d'une mémoire locale. Il existe également des machines à *architecture hybride* qui sont des multi ordinateurs de multiprocesseurs.

4.5.2.1 Machines à mémoire partagée

Les architectures parallèles à mémoire partagée présentent une mémoire contiguë unique accessible directement par leur ensemble de processeurs via un bus ou une hiérarchie de bus (figure 4.4). La plupart des machines à mémoire partagée sont des systèmes symétriques (SMP, Symetric MultiProcessing) où tous les processeurs ont les mêmes fonctions et se disputent les ressources du système de façon uniforme.

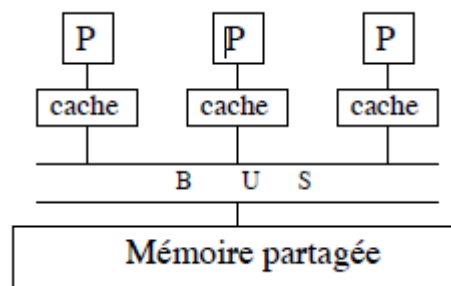


Figure 4.4. Machine multiprocesseurs à mémoire partagée. p représente un processeur indépendant.

L'existence d'une mémoire commune entre les processeurs simplifie le travail de parallélisation des algorithmes. La communication entre les processeurs est facilitée par le fait que les communications se font par le biais de lectures et d'écritures de zones de mémoire partagée. Ceci rapproche la logique de construction de programmes de celle de la programmation séquentielle traditionnelle. D'un autre côté, l'existence d'une mémoire commune est également le point faible de ce type d'architecture en constituant un goulot d'étranglement car tous les processeurs doivent y accéder.

Dans le cas idéalisé du modèle PRAM (Parallel Random Access Machine), souvent utilisé pour étudier d'un point de vue théorique des algorithmes parallèles [81, 82], tout processeur accède en une même quantité de temps tout élément de la mémoire. En pratique, le fait d'augmenter l'architecture introduit généralement une forme de hiérarchie mémoire. En

particulier, la contention d'accès à la mémoire partagée peut être réduite en conservant dans le cache de chaque processeur des copies des données fréquemment accédées. En effet, les accès dans les caches des processeurs sont beaucoup plus rapides que ceux effectués dans la mémoire partagée, mais apparaissent alors des problèmes de cohérence et de synchronisation. Mais malgré l'emploi de plusieurs niveaux de mémoires caches, ces architectures sont souvent limitées en mémoire d'exécution. Elles ne sont donc pas adaptées au parallélisme à très grande échelle en tant que machine unique.

4.5.2.2 Machines à mémoire distribuée

Une machine parallèle à mémoire distribuée comporte un ensemble de nœuds reliés par un réseau de communication. Chaque nœud est composé d'un processeur et d'une mémoire locale. La mémoire est ainsi distribuée parmi les processeurs et n'a donc pas de localisation centrale. Toutes les interactions entre les nœuds doivent passer obligatoirement par le réseau. La figure 4.5 illustre ce type de modèle d'architecture parallèle.

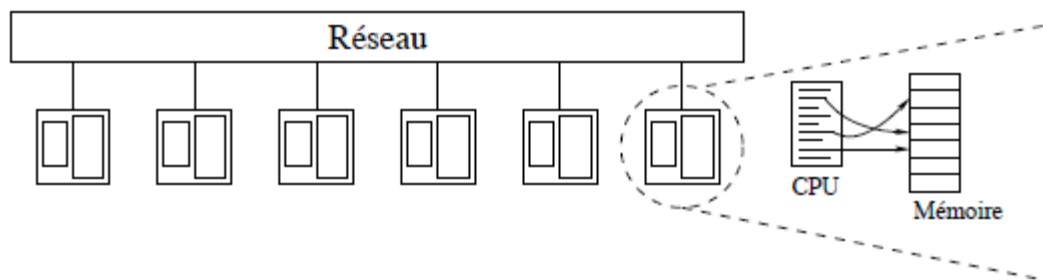


Figure 4.5 Modélisation d'une machine parallèle MIMD à mémoire distribuée. Chaque nœud est composé d'un processeur et d'une mémoire locale. Les nœuds communiquent entre eux en envoyant et en recevant des messages via le réseau de communication.

Ce modèle d'architecture présente un asynchronisme physique dont il faut tenir compte lors de la programmation, provenant du fait que chaque processeur travaille indépendamment avec sa propre mémoire locale.

De telles architectures peuvent atteindre une taille beaucoup plus importante en terme de nombre de processeurs et de capacité mémoire que celles à mémoire partagée. En effet il n'y a pas de contention entre les processeurs pour accéder à la mémoire, la seule contrainte d'extension étant la capacité du réseau de communication. Ainsi, des machines composées de plus d'une dizaine de milliers de processeurs ont vu le jour. Ce sont les architectures dites massivement parallèles (MPP, Massively Parallel Processing).

4.5.2.3 Machines graphiques parallèles

En informatique graphique, les machines Silicon Graphics Incorporated (SGI) Origin et Onyx représentent les standards en terme de puissance. Les SGI Origin 2000 [83] sont basées à la fois sur une mémoire partagée et sur un système distribué. Le système est appelé Scalable Shared-Memory Processor (SSMP). Deux processeurs sont connectés via un “hub” à une mémoire partagée pour former un *nœud*. Plusieurs nœuds sont ensuite connectés en utilisant un réseau d’interconnexion afin de former le système Origin. La figure 4.6 illustre cette architecture.

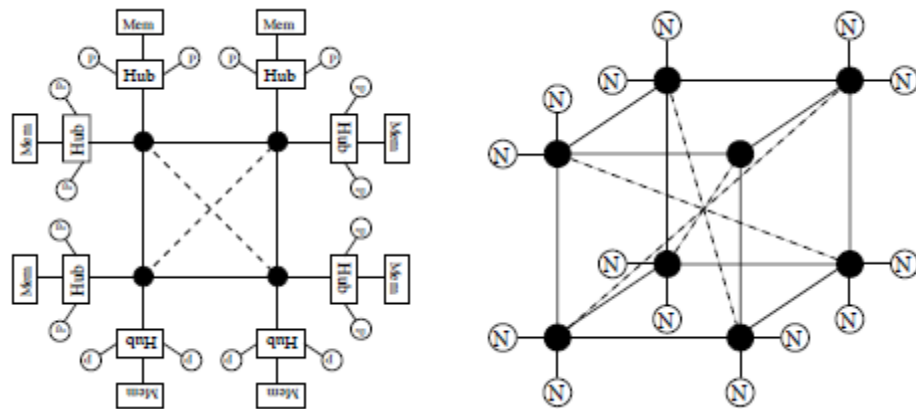


Figure 4.6 Architecture du système SGI Origin. Deux processeurs sont connectés via un « hub » à une mémoire partagée pour former un « nœud ». Plusieurs nœuds sont ensuite connectés en utilisant un réseau d’interconnexion pour former le système global.

Au sein de ce système, la mémoire distribuée est vue comme un unique espace mémoire, tout processeur pouvant accéder à tout emplacement de la mémoire. La dernière configuration de cette machine, la SGI Origin 3000, comporte 512 processeurs.

4.5.2.4 Grappes de machines ou clusters

Depuis quelques années, les *grappes de machines* ou cluster [84] ont fait leur apparition. Une grappe peut se limiter à une grappe d’ordinateurs personnels standards interconnectés par un réseau haut débit (à gauche de la figure 4.7), mais peut également consister en un système de plusieurs machines SMP inter-connectées via un bus de communications et d’entrées/sorties à hautes performances (à droite de la figure 4.7). Ce type d’architecture est alors dite *hybride*, car elle combine les deux types d’architectures présentées précédemment. Avec cette configuration, nous avons à la fois la mémoire partagée au sein de chaque nœud et la mémoire distribuée entre l’ensemble des nœuds. C’est-à-dire qu’elle profite

à la fois de la facilité d'accès aux données offerte par les mémoires partagées, tout en gardant l'extensibilité des machines à mémoire distribuée.



Figure 4.7 (A gauche) Grappe de PCs du projet iCluster (HP, INRIA, ID-IMAG) composée de 226 mono-processeurs Pentium III à 733 MHz avec 256 Mo de mémoire, inter-connectés via un réseau de 100 Mbit/s . (A droite) Grappe du projet iCluster2 composée de 104 bi-processeurs Itanium-2, 64 bits à 900 MHz avec 3 Go de mémoire, inter-connectés par un réseau Myrinet.

Cette technologie des grappes s'oppose à celle plus coûteuse des multiprocesseurs en permettant l'obtention à un *moindre coût* d'une machine parallèle ayant un grand nombre de processeurs avec une mémoire d'exécution importante. Ce nombre de processeurs et la taille de cette mémoire étant limités dans les machines multiprocesseurs (512 processeurs pour la SGI Onyx 3000). De plus une grappe a l'avantage d'être *évolutive*. En effet dans une grappe, la puissance globale de calcul peut être augmentée graduellement en ajoutant un autre système standard.

4.5.3 Parallélisation du processus de traitement

D'après la motivation ci-dessus, notre objectif est de partager le traitement sur plusieurs ordinateurs reliés pour former un réseau informatique, cette architecture parallèle utilisée s'appelle *MIMD* (Multiple Instructions multiple data) Multiple Instructions multiple data comme indiqué le schéma suivant (figure 4.8):

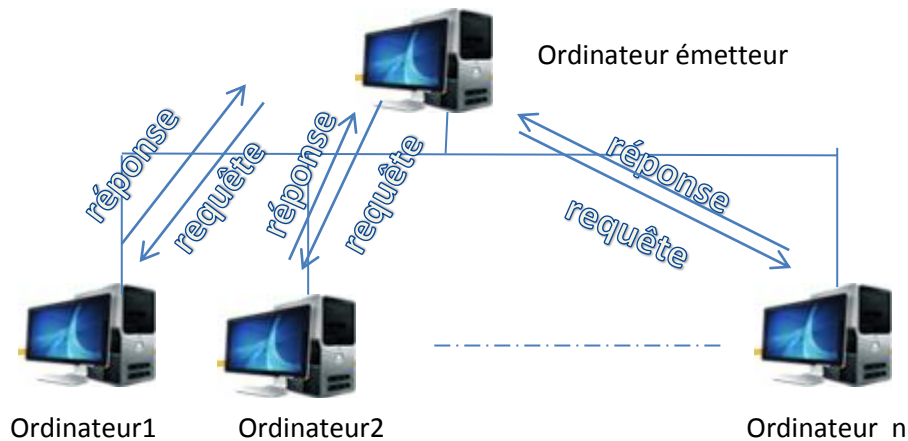


Figure 4.8 Schéma du réseau informatique

Dans la phase online, la comparaison entre l'arbre requête et l'arbre de référence en cours consiste à calculer la distance euclidienne entre chaque 2 nœuds homologues dans tous les niveaux au total, nous avons fait le même calcul de distance 21 fois.

Chaque niveau sera traité sur une machine physique (ordinateur) indépendante. D'après la figure 4.8, on a besoin de 21 machines ($n=21$), même si on travaille sur 3 machines ($n=3$) par exemple, le temps d'exécution sera minimisé. L'ordinateur émetteur envoie la tâche de calcul de distance euclidienne δ_i ($i=2-21$) à tous les ordinateurs sous forme de requête, en même temps il calcule la distance euclidienne δ_1 du premier niveau entre les deux nœuds racines homologues, chaque ordinateur fait son calcul de distance puis envoie le résultat de calcul à l'ordinateur émetteur sous forme de réponse. Ce traitement peut être modélisé comme suit :

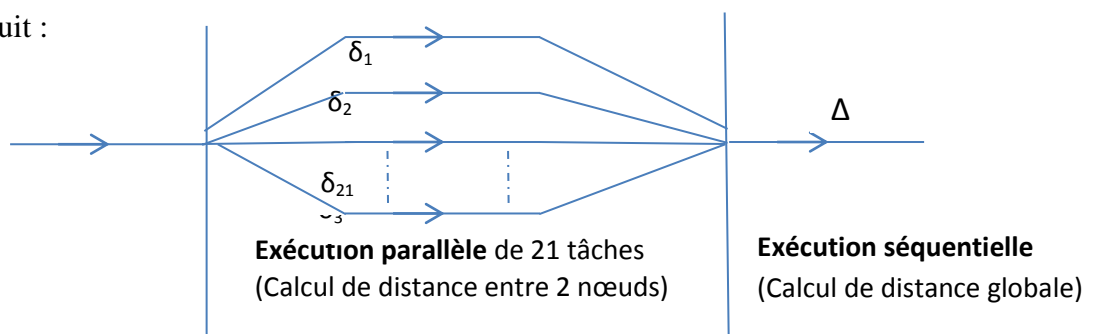


Figure 4.9 modélisation de notre proposition de parallélisme

4.5.3.1 Architecture matérielle

Notre approche est destinée à être exécutée sur une architecture MIMD à mémoire distribuée et plus précisément sur un réseau de station de travail. Ces stations sont monoprocesseur. Chaque nœuds est un PC de type Intel Core i5 (2.40 Ghz, 2.40 Ghz) et il dispose d'une mémoire vive de taille 4 GO.

4.5.3.2 Modèle d'exécution

La parallélisation du programme d'annotation automatique d'image par portion (quadrant) d'image sur architecture MIMD à mémoire distribuée convient bien avec une organisation Maître/Esclave. Le premier pas du processus parallèle d'annotation automatique d'image est de diviser l'image requête en 4 quadrants afin de construire l'arbre quaternaire équilibré à 3 niveaux ce traitement est faite par l'ordinateur émetteur (Maître). Chaque quadrant de l'image est associé à un PC client (Esclave). Chaque client permet d'effectuer un traitement autonome. Il permet de l'extraction du vecteur descripteur adéquat du quadrant, puis effectue la comparaison entre ce vecteur descripteur et le vecteur descripteur du quadrant homologue d'image de référence en cours pour trouver la distance euclidienne d'une manière totalement indépendante des autres clients. Ces derniers renvoient les résultats de calcul au maître qui consiste à calculer la distance globale Δ et le compare avec un seuil donné (après la réception de tous les résultats). Cette tâche est effectuée séquentiellement.

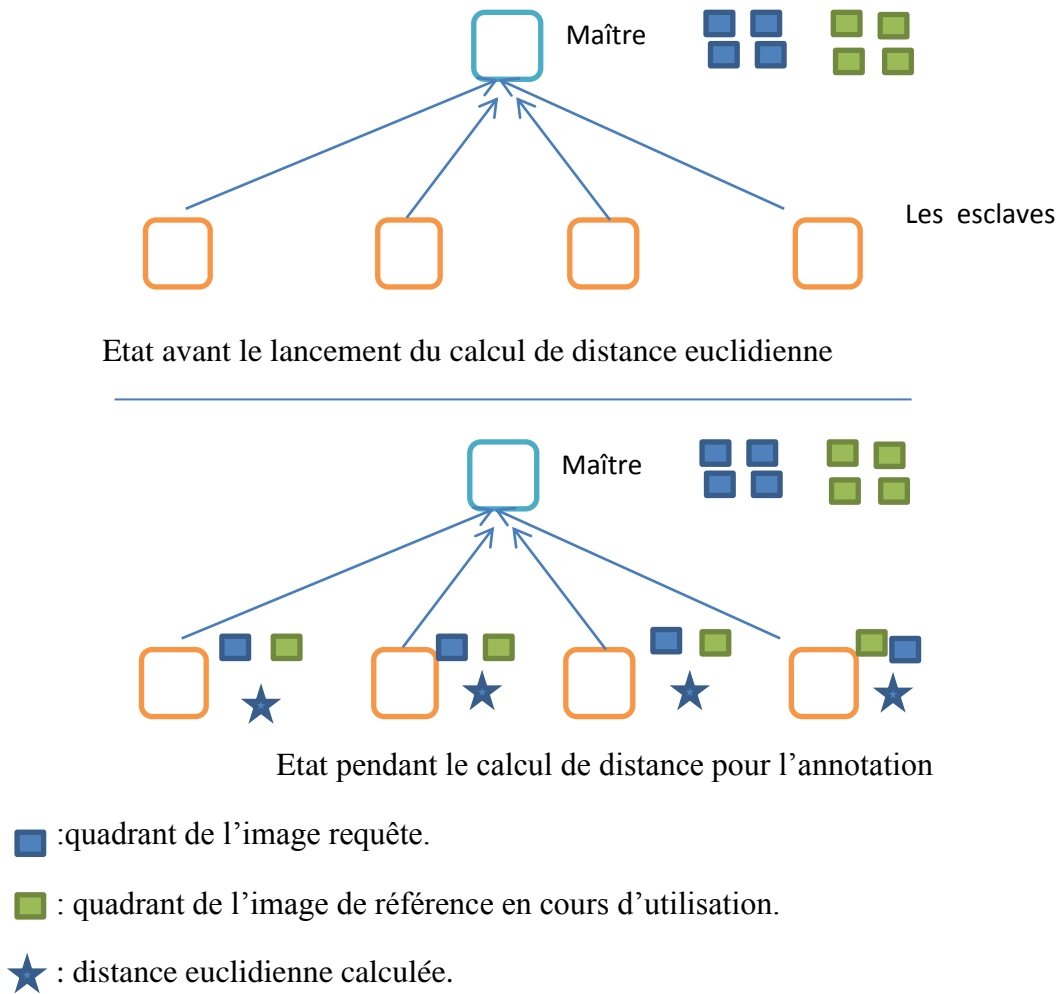


Figure 4.10 Modèle d'exécution Maître/Esclave

4.5.3.3 Fonctionnement de l'algorithme

Donc on peut conclure d'après ce contexte qu'on a 2 procédures exécutées sur cette architecture. Une procédure client (esclave) est installée sur chaque machine du réseau c'est la 1^{ère} procédure (*Procedure CalcDistEucl*). Une procédure maître est installée sur une machine, c'est la 2^{ème} procédure (*Procedure calculdist*).

La 1^{ère} procédure : Calculer la distance euclidienne entre 2 nœuds homologues.

Procedure CalcDistEucl(Nreq,Nref)

Debut

V1 : vecteur descripteur Nreq ;

V2 : vecteur descripteur Nref ;

$DistEucl(V1, V2) = \left(\sum_{k=1}^n |V1(K)_{requete} - V2(K)_{courant}|^2 \right)^{1/2}$

Fin.

La 2eme procédure : calculer la distance entre 2 arbres quaternaires

Cette procédure utilise la 1ere procédure CalcDistEucl(Nreq,Nref) pour tous les 21 nœuds homologues.

Procédure calculdist(AQ_{req}, AQ_{ref})

Debut

Pour i de 1 à 21 faire

CalcDistEucl(Nreq,Nref) ;

$$Distance_{total}(AQ_{req}, AQ_{ref}) = \Delta(i,j) = \frac{\sum_k c_k DistEucl(V1,V2)}{\sum c_k}$$

Fin.

Notations

Nreq : nœud de l'arbre quaternaire qui représente l'image requête

Nref : nœud de l'arbre quaternaire qui représente l'image en cours de la base de référence.

AQ_{req} : l'arbre quaternaire qui représente l'image requête.

AQ_{ref} : l'arbre quaternaire qui représente l'image en cours de la base de référence.

4.5.3.4 Algorithme d'annotation parallèle

Maître (host) :

Etablir la connexion avec les clients.

Charger l'image requête.

Charger l'image de la base de référence en cours.

Envoyer les données initiales (quadrant de l'image) aux clients.

Terminer=faux ;

Tant Que non (Terminer) Faire

Bloquer jusqu'à un client envoie un signal de fin de calcul.

Marquer le client i comme terminer ;

Nbt=Nbt+1 ;

Si Nbt=Nombre de clients Alors

Terminer=vrai ;

FinSi ;

Fin Tant Que ;

Pour chaque client Faire

Collecter les résultats à partir du client.

Fin Pour ;

Exécuter la procédure calculdist(AQ_{req}, AQ_{ref}) ;

Si Δ < seuil Alors Annoter l'image requête par les mots clés sockés dans le nœuds racine ;

FIN.

Client (Node) :

Recevoir les données initiales (les deux quadrants homologues) ;

Calculer le vecteur descripteur du quadrant de l'image requête ;

Exécuter la procédure CalcDistEucl(Nreq,Nref) ;

Envoyer les résultats obtenus vers le maître ;

FIN.

4.6 Conclusion

Ce chapitre présente la conception d'un nouveau système d'annotation automatique d'image rétinienne par le contenu basée sur la structure arborescente afin de faciliter et diagnostic de la rétinopathie diabétique (cas micro anévrisme).

Notre approche d'annotation est basée sur les descripteurs visuels de MPEG7 et l'arbre quaternaire équilibré à trois niveaux. Ensuite nous avons introduire la parallélisation dans le processus d'annotation automatique de notre approche pour gagner du temps d'exécution en utilisant le réseau informatique de plusieurs machines connectées entre-deux. Dans ce qui suit, nous montrons les résultats de l'annotation automatique d'image afin de valider notre approche.

CHAPITRE 5

IMPLEMENTATION ET TESTS

5.1 Introduction

L'objectif de ce chapitre est de présenter une vue globale sur la mise en place de notre technique d'annotation automatique d'images rétiniennes. Il illustre les résultats d'expérimentation par l'algorithme présenté dans le chapitre précédent.

Dans ce chapitre, nous allons d'abord décrire l'implémentation de notre approche proposée à travers un système logiciel d'annotation automatique. Par la suite, nous réalisons un prototype afin de tester notre système sur une série de 60 images issues des bases de données Messidor. Ensuite, nous allons faire une simulation en cas de parallélisme.

5.2 Description de la base MESSIDOR

Les 1200 images couleur de fond d'œil pour la base d'images Messidor ont été acquises par 3 centres d'ophtalmologie en utilisant une caméra vidéo couleur 3CCD sur un rétinographe Topcon TRC NW6 non-mydratique avec un angle de vue de 45°. Les images sont capturées en utilisant 8 bits par canal couleur à la résolution 1440*960, 2240*1488 ou 2304*1536 pixels [85].

800 images ont été acquises avec dilatation de la pupille (une goutte de Tropicamide à 0,5%) et 400 sans dilatation.

Les 1200 images sont découpées en 3 ensembles correspondant à chaque centre d'ophtalmologie. Chaque ensemble est divisé en 4 sous-ensembles zipés contenant chacun 100 images au format TIFF et un fichier Excel avec le diagnostic médical pour chaque image.

5.3 Implémentation de l'arbre quaternaire

Notre avons programmé notre application avec le langage C# de Framework .net version 3.5, et pour l'implémentation de l'arbre quaternaire nous avons créé deux classe

quadtreeNode (le nœud de l'arbre quaternaire) et classe QuadTree qui représente l'arbre quaternaire.

Toutes les méthodes et les propriétés de ces classes nous détaillons maintenant :

5.3.1 Class QuadtreeNode

5.3.1.1 Les attributs

- **Position** : énumération contient les positions de nœud dans l'arbre a comme valeurs {NW, NE, SW, SE, ROOT (pour le nœud racine)}
- **ID** : c'est l'identifiant de nœud.
- **Depth** : c'est la profondeur de nœud.
- **Poids** : c'est le poids de nœud.
- **Parent** : c'est le parent de nœud (parent=null si le nœud est racine).
- **Fils** : collection des objets quadtreeNode contient les quatre fils d'un nœud.
- **Bound** : rectangle représente les bords de nœud.
- **Image** : image représente une partie de l'image de l'arbre quaternaire.
- **Vector** : un tableau représente le descripteur visuel sauvegardé dans le nœud.

5.3.1.2 Les méthodes

- **quadtreeNode ()** : constructeur par défaut de nœud
- **quadtreeNode (parent, position, rectangle, image, x, y, x1, y1)** : le constructeur principale de nœud.
- **subTree (nœud)** : méthodes chargé de deviser un nœud en et retourner les quatre fils.
- **Extractor ()** : méthode chargé de extraire les descripteurs visuel dans le nœud.

5.3.2 Class QuadTree

5.3.2.1 Les attributs

- **root** : racine de l'arbre quaternaire de type quadtreeNode.

5.3.2.2 Les méthodes

- **QuadTree ()** : constructeur par défaut de l'arbre quaternaire.
- **QuadTree (image)** : constructeur charge de crée l'arbre quaternaire.
- **getAllChild ()** : méthode qui retourne tous fils de l'arbre quaternaire.

- getChidNode** (nœud) : méthode qui retourne tous fils de d'un nœud d'arbre quaternaire.
- countInDepth** (profondeur) : méthode qui compter le nombre des nœuds d'un niveau.
- getNodeInDepth** (profondeur) : méthode qui retourne les nœuds d'une profondeur.
- getAllImage** () : méthode qui retourne tous les images des nœuds d'arbre quaternaire.
- getImageInDepth** (profondeur) : méthode qui retourne tous les images des nœuds d'arbre quaternaire pour une profondeur donné.
- getSubTree**(ID) : méthode qui retourne tous les fils d'un nœud à partir de son ID.
- countInSubtree**(ID) : méthode qui compter le nombre des fils (nœuds) dans un sous arbre d'un nœud à partir de son ID.
- distance** (QuadTree, QuadTree) : méthode qui calcule distance entre deux arbres quaternaires.

5.4 Experimentations et Resultats

Dans le cadre de ce travail, nous avons traiter notre approche sur 60 images rétiniennes (dont 40 images pour une base d'images d'apprentissage et 20 images pour une base d'images de test (images requêtes)) telque la base d'apprentissage contient 20 images saines et 20 images pathologiques.

L'ensemble des 60 images sélectionnées est un sous l'ensemble d'images de la base MESSIDOR [86] du fait qu'on s'intéresse à la détection que les images rétinienne saines et celles qui sont infectées par des microanivrismes, la figure suivante présente deux exemples d'images rétiniennes :



Figure 5.1 Exemple d'images rétiniennes de la base de données MESSIDOR ;

(a) : Image saine (normale) ; (b) : Image pathologique

Dans le système d'annotation automatique d'images, l'utilisateur s'intéresse aux reponses pertinentes du système.

Deux critères bien connus pour l'évaluation d'un tel système d'annotation pour mesurer leur performance. Il s'agit des scores suivants :

$$Précision = \frac{\text{nombre d'images correctement annotées}}{\text{nombre d'images annotées par notre algo}} \quad (5.1)$$

$$Rappel = \frac{\text{nombre d'images correctement annotées}}{\text{nombre d'images d'apprentissage}} \quad (5.2)$$

Pour chaque type d'images de la base de test (saines et pathologiques), nous avons déterminé le nombre d'images annotées par notre algorithme, le nombre d'images correctement annotées. L'ensemble des tests a été réalisé sur un PC sous Windows 7 disposant d'un processeur Intel Core I5 CPU @ 2.40Ghz, 2.40 Ghz avec 4 Go de RAM et 750 Disque Dur. Le tableau suivant (cf. Tableau 5.1) présente les résultats obtenus de l'application de notre système d'annotation avec les seuils : 15, 18, 20, 25, 30 et 40.

Seuil	Type d'image	Rappel	Précision
40	Image saine	53%	53%
	Avec microanévrismes	53%	53%
30	Image saine	100%	60%
	Avec microanévrismes	100%	50%
25	Image saine	91%	55%
	Avec microanévrismes	90%	50%
20	Image saine	48%	50%
	Avec microanévrismes	45%	47%
18	Image saine	16%	39%
	Avec microanévrismes	15%	38%
15	Image saine	11%	68%
	Avec microanévrismes	10%	67%

Tableau 5.1. Résultats de l'évaluation des images

Les résultats de recherche pour l'image requête pathologique (contient des microanévrismes) n1.jpg qui se trouve a gauche et l'utilisation des descripteurs : Contour (EDH), Texture (EHD) et Couleur (SCD) respectivement dans les niveaux 0, 1 et 2 de l'arbre quaternaire avec le seuil 30 choisi, sont représentés par la figure 5.2.

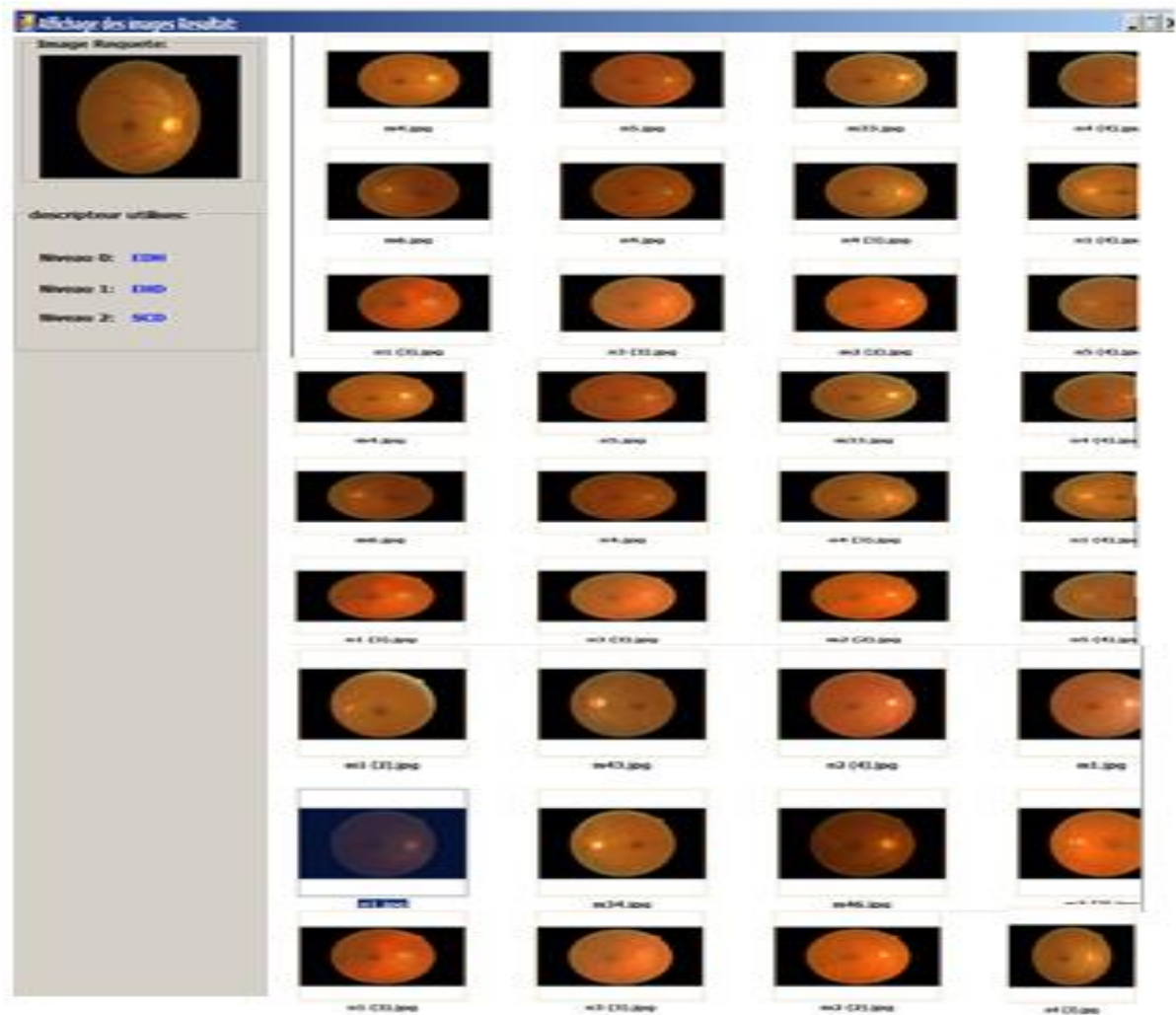


Figure. 5.2 Résultats de recherche pour l'image requête pathologique (Avec micro-anévrisme)

5.4.1 Calcul du temps d'exécution (cas séquentiel)

Dans ce cas, notre approche est implémentée sur une seule machine en se basant sur le calcul du temps d'exécution alloué pour l'extraction des vecteurs descripteurs des quadrants d'image et le calcul des distances euclidiennes entre chaque deux nœuds homologues. Ce temps est présenté par la formule suivante :

$$\text{TEGS} = \text{TEVA} + (\text{TED}_{\text{con}} + \text{TED}_{\text{text}} * 4 + \text{TED}_{\text{coul}} * 16); \quad (5.3)$$

$$\text{TEVA} = \text{TEX}_{\text{con}} + (\text{TEX}_{\text{text}} * 4) + (\text{TEX}_{\text{coul}} * 16); \quad (5.4)$$

TEGS: temps d'exécution général séquentiel entre deux arbres en seconde.

TVA : temps d'exécution pour l'extraction des descripteurs de tous les nœuds de l'arbre.

TEX_{con} : temps d'exécution pour l'extraction du vecteur descripteur pour le contour.

TEX_{text} : temps d'exécution pour l'extraction du vecteur descripteur pour la texture.

TEX_{coul} : temps d'exécution pour l'extraction du vecteur descripteur pour la couleur.

TED_{con} : temps d'exécution pour le calcul de distance euclidienne entre deux vecteurs contour

TED_{text} : temps d'exécution pour le calcul de distance euclidienne entre deux vecteurs texture

TED_{coul} : temps d'exécution pour le calcul de distance euclidienne entre deux vecteurs couleur

Le tableau suivant (tableau 5.2) présente les résultats obtenues pour le TEGS calculé par la formule 5.3, en appliquant le seuil 30 sur une seule machine.

seuil	Nombre D'images	Contour		Texture		Couleur		TEGS (s)
		TEX _{con}	TED _{con}	TEX _{text}	TED _{text}	TEX _{coul}	TED _{coul}	
30	60	0.2	0.07	0.15	0.03	0.19	0.06	4.99

Tableau 5.2 Temps d'exécution général séquentiel entre deux arbres

5.4.2 Calcul du temps d'exécution dans le cas parallèle

Notre approche est implémentée sur 21 machines dont une est un serveur (maître) et les autres sont des clients, alors chaque machine est responsable de faire l'extraction du vecteur descripteur associé au quadrant d'image requête envoyé par le serveur et le calcul de la distance euclidienne entre deux nœuds homologues, pour cela le temps d'exécution général en cas de parallélisation noté par TEGP est présenté par la formule suivante :

$$TEGP = \text{MAX} (T_{Mi}, i=1..21) + TE_{\Delta} + TE_{\text{trasf}} \quad (5.5)$$

TEGS: temps d'exécution général parallèle entre deux arbres en seconde.

$T_{Mi} = TEX_{\text{con}} + TED_{\text{con}}$. Pour le contour.

$T_{Mi} = TEX_{\text{text}} + TED_{\text{text}}$. Pour la texture

$T_{Mi} = TEX_{\text{coul}} + TED_{\text{coul}}$. Pour le couleur.

TE_{Δ} = temps d'exécution pour le calcul de la distance globale Δ entre les deux arbres.

TE_{transf} =le temps de transfert de requête ou de réponse.

Le tableau suivant (5.3) présente les résultats obtenus pour le calcul de TEGP en appliquant le seuil 30 sur l'architecture client/serveur.

Seuil	Nombre d'images	Contour		Texture		Couleur		TE_{Δ}	TE_{transf}	TEGP (s)
		TEX_{con}	TED_{con}	TEX_{text}	TED_{text}	TEX_{coul}	TED_{coul}			
30	60	0.2	0.07	0.15	0.03	0.19	0.06	≈ 0	≈ 0	0.27

Tableau 5.3 Temps d'exécution dans le cas parallèle

5.5. Discussion

A partir du tableau. 5.1, nous remarquons que notre système (le cas séquentiel) permet d'obtenir des résultats intéressants en appliquant un seuil égal à 30. Les résultats sont de moindre qualité quand nous considérons des seuils correspondants à 15,18,20 et 25. Ceci est dû au calcul de la distance associée pour chaque image représentée par l'arbre quaternaire ; comme les microanivrismes sont caractérisés par des points (ou petits cercles), assez contrastes, avec des bords nets, l'exploitation des descripteurs de la forme et de la

texture influent grandement sur les résultats obtenus. D'autre part, la sauvegarde des trois (3) descripteurs dans les différents niveaux (descripteur de contour, EHD (texture) et SCD (couleur) respectivement dans les niveaux de l'arbre quaternaire 1,2 et 3) affecte les résultats donnés par notre système. En contrepartie, cette approche est très couteuse en termes de temps d'exécution comme indiqué dans le tableau 5.2.

Pour cela, nous avons introduire la notion du parallélisme par l'exploitation de l'architecture client/serveur avec 21 machines en tous. Le tableau 5.3 donne le temps d'exécution dans le cas parallèle et en remarque que le temps d'exécution dans ce cas est très petit par rapport au cas séquentiel.

5.6 compromis notre architecture parallèle VS nombre des nœuds

Notre approche de parallélisation s'avère robuste et performante pour faire l'extraction du vecteur descripteur et le calcul de distance euclidienne associé respectivement aux deux nœuds homologues de deux arbres quaternaires. Ceci est dû à l'utilisation d'une architecture client/serveur. Cette dernière est constituée de 21 machines. Le tableau suivant (tableau 5.4) illustre l'évolution du nombre de machines sous réseau par rapport à l'augmentation du niveau de détail en ajoutant le nombre des nœuds.

Nombre des niveaux	Nombre des nœuds	Nombre des machines
3	21	21
4	85	85
5	441	441

Tableau 5.4 Evolution du nombre de machines par rapport au niveau de détail.

On remarque que à partir du niveau 4, notre architecture proposée demande un nombre important de machines, ce qui implique que cette dernière reste adéquate pour 3 niveaux, donc il faut prévoir d'autres architectures de parallélisation pour annoter une image par l'arbre quaternaire plus de 3 niveaux.

5.7. Conclusion

Dans ce chapitre, nous avons présenté l'implémentation d'un prototype de notre système d'annotation automatique d'images et son expérimentation. Cette approche est basée sur l'arbre quaternaire équilibré à trois niveaux et les descripteurs visuels. Les nœuds de chaque niveau (de 0 à 3) stockent un type de descripteur visuel. Ce dernier correspond au contour (obtenu par filtre de sobel), la texture (EHD de MPEG-7) et la couleur (SCD de MPEG-7).

Les résultats obtenus montrent l'efficacité de notre approche de combinaison, et la pertinence des descripteurs de la norme MPEG 7 dans l'arbre quaternaire. Le meilleur score obtenu est de la combinaison de *contours/texture/couleur*,. Ceci est dû à la richesse de caractérisation visuelle en utilisant l'approche de combinaison dans l'arbre quaternaire (global et local), qui a montré une pertinence excellente tout dépend des caractéristiques d'une telle base de données.

Nous avons calculé le temps d'exécution en cas séquentiel et parallèle pour notre approche d'annotation automatique d'images et on a remarqué la grande différence entre les deux temps.

CONCLUSION GENERALE ET PERSPECTIVES

L'objectif principal de notre travail porte sur la parallélisation du processus d'annotation automatique d'images dans le but de développer d'un outil qui servira à une aide au diagnostic en ophtalmologie, notamment pour le dépistage de la Rétinopathie Diabétique, une maladie particulièrement sévère et répandue. La technique d'annotation automatique d'images proposée est basée sur les descripteurs visuels de MPEG-7 et l'arbre quaternaire équilibré à trois niveaux. Cette dernière est utilisée comme structure de données pour indexer les images. Cette indexation consiste à sauvegarder les descripteurs visuels à différents niveaux de l'arbre quaternaire ainsi les mots clés indiquant la pathologie de l'image (micro-anévrismes ou normale) dans le niveau 0. Une distance de similarité entre arbres quaternaires (arbre quaternaire associé à l'image requête et arbre quaternaire en cours de la base d'apprentissage) est calculée, puis comparée à un seuil donné afin d'annoter les images similaires.

Le plan de travail de ce mémoire est consacré à décrire la vision humaine/vision assistée par ordinateur, les principaux traitements d'images et le cadre biomédical du projet (rétinopathie diabétique). Ensuite, nous avons rappelé un état de l'art des différentes approches de recherche d'images connus dans la littérature. Ensuite, nous avons fait un survol sur les modèles et les méthodes déjà utilisées dans l'annotation d'images, puis, nous avons décrit notre approche adoptée pour l'annotation automatique d'images dans le cas séquentiel et parallèle.

L'inconvénient de l'application de cette nouvelle approche d'annotation automatique d'images c'est le temps d'exécution élevé, pour cela cette approche a été parallélisée dont l'objectif est l'amélioration du temps d'exécution du calcul.

Pour résoudre ce problème de temps d'exécution, nous avons proposé l'architecture parallèle MIMD mémoire distribuée qui a été simulée par un réseau de station de travail avec le modèle « client/serveur ». Tous les ordinateurs du réseau sont monoprocesseurs. Chaque poste est un PC de type Intel Pentium ore I5 (2.4 GHZ) avec une mémoire vive de taille 4GO.

Nous remarquons que notre système (le cas séquentiel) permet d'obtenir des résultats intéressants en appliquant un seuil égal à 30. Les résultats que nous avons obtenus dans le cas parallèle (en termes de temps d'exécution) aussi montrent l'efficacité de la méthode parallèle proposée par rapport à la méthode séquentielle.

Plusieurs perspectives peuvent être envisagées dans le prolongement de ce travail de mémoire.

- Le système présenté dans ce mémoire peut également être étendu pour détecter d'autres maladies liées à la RD tels que les néovaisseaux et les nodules cotonneux et d'autres pathologies qui affectent la rétine à savoir : le glaucome, la cataracte.
- Faire la classification de la RD afin de faciliter aux ophtalmologues l'identification du stade de la RD, les Stades de rétinopathie sont :
- Exploitation d'autres distances de calcul pour la comparaison entres arbres.
- Exploitation d'autres descripteurs de MPEG-7 tels que Dominant Color Descriptor (DCD), Homogeneous Texture Descriptor (HTD).
- Nous proposons d'utiliser un réseau de station de travail qui utilise un support de communication de hautes performance comme FDDI ou d'autres architectures parallèles comme les GPUs (Graphic Processing Unit).

REFERENCES BIBLIOGRAPHIQUE

1. Guy almouzni. " Traitement numériques des images 2012/2013 ".école internationale des sciences du traitement de l'information(EISTI)..
2. Gastellu-Etchegorry. " Acquisition et traitement d'image numérique ". Avril 2008 - Université Paul Sabatier IUT - Département de Mesures Physiques J.P..
- 3 Association de la langue française pour l'étude de diabète et des maladies métaboliques. Recommandations pour le dépistage et la surveillance de la rétinopathie diabétique [en ligne]. 1996. page109
- 4.Viswanath K, Murray Mc Gavin DD."Diabetic Retinopathie : clinical finding and management". Communty Eye Health 2003 ; 51 :1776-82.
5. Item 233, "Rétinopathie diabétique (RD) : support de cours 2010/2011".Collège des Ophthalmologistes Universitaires de France.
6. Z. M. da Silva Correa and R. E. Jr, "Aspectos patologicos da retinopatia diabetica,"Revista de Oftalmologia de Sao Paulo, vol. 68, no. 3, pp. 410-414, 2005.
7. R. Chbeir,. "Modélisation de la description d'images : Application au domaine médicale". Thèse de doctorat, Institut national des sciences appliquées de LYON, 2001.
8. Le Thi lan. " Indexation et recherche d'images par le contenu". Mémoire de master, Institut polytechnique de Hanoi, 2005.
9. Trayser G. " Interactive System for Image Selection". Digital Imaging Unit Center of Medical Informatics University Hospital of Geneva Disponible sur: <http://ww.expasy.ch/UIN/html1/projects/isis/isis.html>, 1994.
10. Samet H. "The design and Analysis of Spatial Data Structure". Reading: Addison-Wesley, 1989.
11. Orenstien J. "Spatial query processing in an object-oriented database system". SIGMOD-Record, 1986.

12. Guttman A. "R- trees: a dynamic index structure for spatial searching". SIGMOD Reeced, 1984.
13. Finkel R.A., Bentley J.L. "Quad Trees: A Data structure for Retrieval on Composite Keys". Acta Informatica, 1974.
14. Bentley J.L. "Multidimensional binary search tree used for associative searching". Communications of the ACM, 1975.
15. J. Landre. "Analyse multi résolution pour la recherche et l'indexation d'images par le contenu dans les bases de données". Thèse, Université de Bourgogne (U.F.R), 2005.
16. K. Houari, "Recherche d'image par le contenu". Thèse De Doctorat en Sciences en Informatique. Juin 2010 Université Mentouri Constantine.
17. C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz. "Efficient and effective querying by image content". Journal of Intelligent Information Systems, 3(3/4) :231–262, 1994.
18. S. Lafon, R. R. Coifman, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S., W. "Zucker.Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps". 7426–7431, PNAS, May 24., Vol. 102, no. 21, 2005.
19. S. Lafon, Y. Keller and Ronald R. Coifman. "Data Fusion and Multi-Cue Data Matching by Diffusion Maps". In IEEE transactions on Pattern Analysis and Machine Intelligence, 2006.
20. A. Iannino et S. D. Shapiro. "A survey of the Hough transform and its extension for curve detection. Proc". IEEE PRIP'78, Chicago, mai juin, p. 32-38. 1978.
21. M. Hu. "Visual pattern recognition by moment invariants". IRE Transactions on informatio theory, 8:351–326, 1962.
22. Xiang Sean Zhou, Sonja Zillner, Manuel Moeller. "Semantic and SBIR: A Medical Imaging Perspective". Siemens Medical Solutions, USA, 2008.
23. Niblack W., Xiaoming Zhu, Hafner J.L. "Updates to the QBIC system. Proceeding of the SPIE" – The International Society for Optical Engineering, 1997.
24. Pentland A., Picard R. W., Sclaroff S. "Photobook: tools for contentbased manipulation of image databases". Proceeding of the SPIE International Society for Optical Engineering, 1995.
25. Smith J. R., Chang S. F.. "Querying by color region using the Visual SEEK content-based visual query system. Maybury, MT ed. ", Intelligent Multimedia Information Retrieval, Menlo Park, Calif, 1997.
26. Smith J. R., Chang S. F. "An image and video search engine for the world-Wide_Web". Proceeding of SPIE Storage and Retrieval for Image and Video Database, 1997.

27. Robb R.A., Hanson D.P, "A software System for Interactive and Quantitative Visualization of Multidimensional Biomedical Image". Australasiab Physical and Engineering Sciences in Medicine, 1991.
28. Sensor System,. "MEDx software package", 1999.
29. Deselaers T.. " Features for Image Retrieval". Diploma Thesis, RWTH Aachen University, Aachen, Germany, December 2003
30. Hyvonen E., Styrman A., Saarela Sm,. "Ontology-Based Image Retrieval". <http://www.cs.helsinki.fi/group/seco/presentations/www2003/p199-hyvonen.html>.
31. Mezaris V., Kompatsiaris I., Strintzis M.Gm, " An Ontology Approach to Object-Based Image Retrieval". ICIP 2003.
32. Maillot N., Thonnat M., Boucher A.,. "Towards ontology-based cognitive vision. Machine Vision and Application ", 2004.
33. Cheng S.F, Chen.W., Sundaram H.,. "Semantic Visual templates: linking visual features to semantics". Proceedings of International Conference on Image processing. ICIP'98, Los Alamitos, CA, USA, 1998.
34. Minka T.. An image database browser that learns from user interaction. MIT Media Laboratory, Cambridge MA, MIT Master of Engineering Thesis, 1996.
35. Yamaguchi K., Ohbo N., Kunii Y.L., et al.,. ELF, " Extended relational model for large, flexible picture databases". Proceeding of the Workshop on Picture Data Description and Management, Asilomar, CA, 1980.
36. Hsu C.C., Chu W.W., Taira K.,. " A Knowledge-Based Approach for Retrieving image by content". IEEE Transaction on Knowledge and Data Engineering, 1996.
37. Eakins J.P., Graham M.E. "Content-based Image Retrieval. Technical report of the JISC Technology", Institute for Image Data Research, University of Northumbria at Newcastle, 1999. Disponible sur: <http://www.unn.ac.uk/iidr/research/cbir/report.html>.
38. Baldonado, M., Cousins, S., Gwizdka, J, et Paepeke, A. Notable. " At the Intersection of Annotations and Handheld Technologies". HUC, pp. 100-113. Bristol. 2000.
39. Veron, M. "Modélisation de la composante annotative dans les documents électroniques". Master de recherche, IRIT, Toulouse. P. 53. 1997.
40. Adler. M. et Van Doren. C. "How to read a book. New Yourk. Simon and Schuster. 1972.
41. Damas, L. Mille, A. et Versace, R. "Prendre en compte les comportements cognitifs des apprenants dans la conception de systèmes d'assistance à l'apprentissage humain". TICE 2002,pp.111-118. INSA. Lyon. France. 2002.
42. Huart. P. "Définition d'un poste de lecture active de documents électroniques". Rapport de DEA. Toulouse. IRIT. P. 58 pages. 1996.

43. Bringay. S., Barry. C. et Charlet. J. " les documents et les annotations du dossier patient hospitalier. ", Information- Interaction- Intelligence. 4(1). Pp. 191-2011. 2003.
44. M.Bakalam, S.Oukid, N.Benblidia. "Indexation textuelle sémantique pour un moteur de recherche d'images sur le web". Faculté d'informatique. Université Blida.2010.
45. S. Tollari. " Indexation et recherche d'image par fusion d'informations textuelles et visuelles", thèse de doctorat de l'université du sud Toulon – var (France), Octobre, 2006.
46. Maron O., Ratan A. L., " Multiple-Instance Learning for Natural Scene Classification. ", ICML, 1998, p. 341-349.
47. M. Yasuhide T. H., Ryuichi O., " Image-to-Word transformation Based on dividing and vector quantizing image with words ", In MISRM'99, First international workshop on multimedia intelligent storage and retrieval management, 1999.
48. Duygulu P., Barnard K., DE Freitas J. F. G., Forsyth D. A., " Object Recognition as Machine Translation : Learning a Lexicon for a Fixed Image Vocabulary. ", ECCV (4), 2002, p. 97-112.
49. Jeon J., Lavrenko V., Manmatha R., " Automatic image annotation and retrieval using cross-media relevance models. ", SIGIR, 2003, p. 119-126.
50. Monay F., Gatica-Perez D., "On image auto-annotation with latent space models. ", ACM Multimedia, 2003, p. 275-278.
51. LI J., Wang J. Z., "Automatic Linguistic Indexing of Pictures by a Statistical Modeling Approach. ", IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no 9, 2003, p. 1075-1088.
52. R. Albatal. "Phrases Visuelles pour l'annotation automatique d'images", Laboratoire d'informatique de Grenoble (LIG).2010
53. H. Bannour, Céline Hudelot . " Construction de hiérarchies sémantiques pour l'annotation d'images". Revue d'Intelligence Artificielle, Lavoisier (Hermes Science Publications), 2013.
54. Le Thi Lan. " Indexation et recherche d'images par le contenu". Mémoire de Master, Institut polytechnique de Hanoi, 2005.
55. Gustavo Carneiro et Nuno Vasconcelos, "Formulating Semantic Image Annotation as a Supervised Learning Problem", 2004
56. J. Jeon, V. Lavrenko and R. Manmatha. Automatic Image Annotation and Retrieval using Cross-Media Relevance Models. Center for Intelligent Information Retrieval, Computer Science Department, University of Massachusetts, 2003.

57. Nguyen huu luong Tu. "Recherche d'image basée sur le contenu sémantique". Travail d'Intérêt Personnel Encadré, Institut de la Francophonie pour l'Informatique, 2005.
58. Sabine Barrat¹. Salvatore Tabbone. "Classification et extension automatique d'annotations d'images en utilisant un réseau Bayésien. LORIA-UMR7503 - Université Nancy 2. 2010
59. Barnard K., Duygulu P., Forsyth D., De Freitas N., Blei D. M., Jordan M. I., Matching words and pictures, *Journal of Machine Learning Research*, vol. 3, no 6, pp. 1107–1135, 2003.
60. Grosky W. I., Zhao R., Negotiating the Semantic Gap : From Feature Maps to Semantic Landscapes, in *SOFSEM '01*, pp. 33–52, 2001.
61. Benitez A., Shih-Fu C., Perceptual knowledge construction from annotated image collections, in *ICME '02*, vol. 1, pp. 189–192, 2002.
62. Hakim Hacid. Annotation Semi-automatique de Grandes BD Images : Approche par Graphes de Voisinage. Laboratoire ERIC - 5, Université Lumière Lyon 2.2006.
63. Zhang, Y. J. (2006). An Overview of Image and Video Segmentation in the Last 40 Years.
64. Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. 1990. Indexing by Latent Semantic Analysis. *Journal of the American Society of Information Science*, 41(6), 391-407.
65. La wCascia, M., Sethi, S., & Sclaroff, S. 1998 (6). Combining Textual and Visual Cues for Content-based Image Retrieval on the World Wide Web. In: *IEEE Workshop on Content-based access of Image and Video Libraries*.
66. F. Monay and D. Gatica-Perez, "On Image Auto-Annotation with Latent Space Models," *Proc. ACM Multimedia*, 2003.
67. F. Monay & D. Gatica-Perez, "PLSA- based image auto annotation: constraining the latent space". In proceedings of ACM Multimedia (ACM MM). New york, NY, USA: ACM Press, 2004.
68. Sabrina Tollari. Image indexing and retrieval by combining textual and visual informations. laboratoire LSIS - UMR CNRS 6168. Thèse de Doctorat de l'Université du Sud Toulon-Var.2006.
69. H. Haddad, P. Mulhem . Utilisation de la Fouille de Données Images pour l'Indexation Automatique des Images. *INFORSID* , Genève, Switzerland. pp.405--418, 2001
70. X-J. Wang, L. Zhang, F. Jing, W-Y. Ma. AnnoSearch: Image Auto-Annotation by Search. *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. New York, June, 2006.

71. SAMET H., « The Quadtree and Related Hierarchical Structures », *Computing Surveys*, vol. 16, n° 2, 1984, p. 187-260.
72. Manouvrier M., Rukoz M., Jomier G., « Quadtree representations for storage and manipulation of clusters of images », *Image and Vision Computing*, vol. 20, n°7, 2002, p. 513-527.
73. Marta Rukoz, Maude Manouvrier, Genevieve Jomier, « Distances de similarite d.images basees sur les arbres quaternaires » 18emes Journees Bases de Donnees Avancees . 21-25 Oct./2002. BDA.2002, pages 1 a 20.
74. Castelli V., « Image Databases : Search and Retrieval of Digital Imagery », chapitre 14 - Multidimensional Indexing Structures for Content-Based Retrieval, p. 373-433, Wiley Inter-Science, 2002, V. Castelli and L.D. Bergman (Eds) - ISBN : 0-471-32116-8.
75. I.Nedjar.« Content Medical Based Image Retrieval ».Mémoire de magister. Université Abou Bakr Belkaid– Tlemcen Faculté des Sciences Département d’Informatique.2012.
76. Antoine COUTANT. La méthode des contours actifs en traitement des images. Mémoire pour l’examen probatoire en Calcul Scientifique. Conservatoire National des Arts et Métiers.2005.
77. M. Sekma, A. Ben abdelali, A. Mtibaa. Application d’un Descripteur MPEG7 de Texture pour la Segmentation Temporelle de la Vidéo. STIT IEEE.2012. Sousse (Tunisia)21-24 march 2012.
78. M.Rukoz, M.Manouvrier,G.Jomier. Distances de similarité d’images basées sur les arbres quaternaires. Coopération internationale CNRS – CONICIT (accords 8680 et 10058). 18èmes Journées Bases de Données Avancées – 21-25 Oct. /2002. BDA’2002, pages 1 à 20.
79. M.-J. Flynn. Somme Computer Organizations and their effectiveness. *IEEE Transactions on Computers*, C-21(9) :948-960, 1972.
80. A. Tanenbaum. *Modern Operating Systems*. Prentice Hall, 1992.
81. J-L. Roch. Complexité parallèle et algorithmique PRAM. In Ecole d’automne Conception et Analyse d’Algorithmes Paralleles (CAPA 93), Port d’Albret, Landes, France, Septembre 1993.
82. G. Authié, J-M. Garcia, A. Ferreira, J-L. Roch, G. Villard, J. Roman, C. Roucairol, and B. Virot. *Parallélisme et applications irrégulières*. Hermès, 1995.
83. M. Grobe. *The architecture and use of the Origin 2000*. Academic Computing Services, University of Kansas, January 1998.
84. T. Sterling, D. Savarese, D. J. Becker, J. E. Dorband, U. A. Ranawake, and C. V. Packer. *Beowulf : A parallel workstation for scientific computation*. In *Proceedings of the 24th*

International Conference on Parallel Processing, pages I :11–14, Oconomowoc, WI, 1995.

85. <http://messidor.crihan.fr/>

86. Feroui. A. Analyse des images couleur du fond d'œil pour l'aide au diagnostic en ophtalmologie: application a la détection des pathologies rétiniennes. Thèse de doctorat, université abou bakr belkaid Tlemcen. Faculté de technologie. mai 2014.