



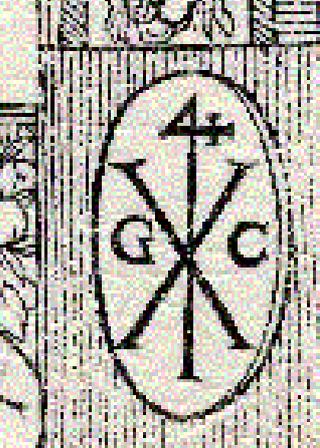
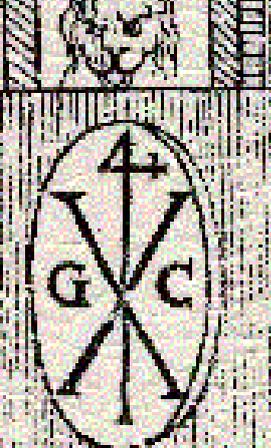
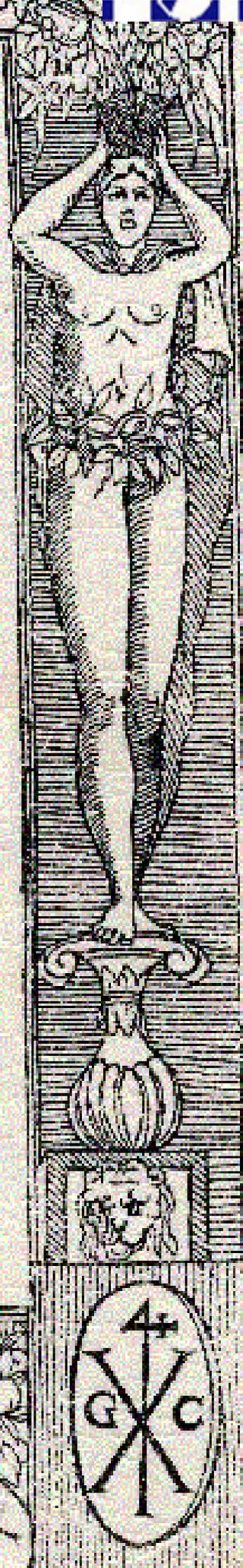
Université La Rochelle
Rapport de Stage
Master 2 Informatique
Mathématiques et Applications
2004-2005

**Extraction
des signatures d'images graphiques
pour l'indexation d'images: applications à
la valorisation du patrimoine**

Etudiant :
Mouhamed Hammoud
Mhammoud@univ-lr.fr

Responsable de stage :
Jean-Marc Ogier
Jean-marc.ogier@univ-lr.fr
Patrick Franco
Patrick.Franco@univ-lr.fr

Laboratoire d'accueil :
Laboratoire Informatique, Image, Interaction (L3i)
17042 La Rochelle Cedex 1 – FRANCE



Résumé

Ce rapport a été réalisé lors de mon stage de fin de Master 2 IMA à l'Université de La Rochelle. Celui-ci s'est déroulé au sein du laboratoire L3I du 2 février au 7 juillet 2005. Le but de ce stage concernait **l'Extraction des signatures d'images graphiques pour l'indexation d'images: applications à la valorisation du patrimoine.**

Ce rapport est donc composé de trois parties :

- Le contexte du stage
- Etat de l'art des différents systèmes de représentation d'images classiques.
- Deux nouvelles méthodes pour extraire les signatures utilisées pour la recherche d'image par le contenu.

Mots clés : recherche des images par le contenu, image graphique, relation spatial, MST, attributs géométriques, signature.

Remerciement

Je tiens à remercier ici mon maître de stage, M. Jean-Marc Ogier, pour m'avoir accueilli au sein du laboratoire L3i, et pour sa grande disponibilité tout au long du stage. Nous avons eu beaucoup de discussions très intéressantes qui m'ont permis d'envisager les choses sous un autre angle.

Je remercie également tous les enseignants-chercheurs, doctorants et stagiaires du L 3i pour leur aide, mes parents qui m'ont soutenu toute cette période ainsi mon frère Amine pour m'avoir permis de faire cette année d'études.

Résumé	2
Remerciement.....	3
Introduction.....	5
Chapitre 1 : Contexte du stage	6
Chapitre 2 : Présentation du Laboratoire L3i.....	7
Chapitre 3 : Indexation et recherche d'images [1]	13
Chapitre 4 : Développement	25
Chapitre 5: Conclusion et perspective.....	35
Bibliographie	37

Introduction

L'indexation des images par le contenu est un objectif ambitieux qui est rendu possible par les capacités techniques offertes par les matériels actuels, et aussi par les progrès scientifiques du traitement d'images et de la vision par ordinateur. Cet objectif se décline en deux familles de problèmes de complexités différentes :

- Etre capable, suite à une indexation, de retrouver une image à partir d'une image pratiquement identique ou simplement à partir d'un seul fragment.
- Etre capable de retrouver des classes similaires, même si les images sont très différentes sur le plan du signal: paysage de montagne, etc.

Les applications des outils dérivant de ces travaux concernent bien sûr la consultation et la recherche dans les bases d'images fixes :

- Consultations en vue de vérification de copyright (presse, industrie textile, contrôle aux frontières),
- Recherche (agence de presse, agence de photos pour publicitaires, bases d'images scientifiques ou industrielles.)

Dans le cadre de ce sujet, il s'agit de mettre en place des techniques de calcul de signature sur des images graphiques, permettant d'indexer les images et de faciliter la recherche d'informations dans ces images.

Ce rapport s'articulera donc autour de trois axes :

Tout d'abord, nous présenterons le contexte du stage, le sujet proposé et les problématiques posées et le laboratoire L3i.

Ensuite, nous nous intéressons plus spécifiquement à l'indexation des images. On traite les différents systèmes de représentation d'images classiques. Nous présentons l'intérêt de la notion de l'approche spatiale pour la description d'image.

Enfin, le reste du rapport portera sur l'ensemble du travail effectué : la méthodologie, les choix effectués, les résultats obtenus, leur analyse et les perspectives envisageables.

Chapitre 1 : Contexte du stage

1.1 Contexte scientifique :

Participant à une action concertée incitative CNRS/INRIA/MENRT, ce projet s'inscrit dans une démarche de sauvegarde et de valorisation de données patrimoniales dont la communauté internationale a pris conscience de l'intérêt, comme en attestent les impulsions prises au niveau européen par les représentants nationaux des grands organismes de gestion du patrimoine. Le contexte de l'étude concerne les collections d'ouvrages numérisés, qui constitueront à très court terme des entrepôts gigantesques de données, représentés sous forme d'images scannées, pour lesquelles les techniques traditionnelles des bases de données sont inopérantes. L'exploitation et la valorisation à venir de ces collections d'images n'ont toujours pas trouvé de réponse satisfaisante, du fait même de leur caractère faiblement structuré. La génération de ces entrepôts de données, présentés sous forme de collections de documents hétérogènes faiblement structurés soulève le problème de la recherche d'information et de la navigation au sein de ces corpus.

1.2 Le but de ce stage :

Il s'agit de déterminer des indices s'adaptant aux différentes représentations de l'information que l'on peut rencontrer dans ces documents patrimoniaux comme des zones textuelles, imprimées ou manuscrites, des images, des illustrations graphiques. L'extraction de signatures sur ces zones apporteront des connaissances spécifiques qui aideront la navigation et la recherche d'informations.

Dans le cas particulier de cette étude, l'objet d'étude concernera principalement l'indexation des lettrines, caractéristiques graphiques très fréquemment utilisées sur les documents anciens. Pour l'étude, il s'agira de tester des jeux de signatures déjà existantes au L3i (attributs couleurs, signatures statistiques, structurelles), afin de mesurer l'opportunité de leur utilisation dans une telle chaîne d'indexation. Une stratégie de recherche sera dès lors établie en fonction des premiers tests opérés : développement de nouvelles signatures, combinaison optimale de vecteurs de caractéristiques, fusion de classifieurs,

Les données sur lesquelles seront testés les algorithmes sont des images de documents anciens issus de bibliothèques du centre d'études supérieures de la renaissance de Tours.

Le stage s'est déroulé sous la tutelle de **Jean-Marc Ogier**, **Patrick Franco** dans le laboratoire L3i. Le développement du programme a été réalisé sous Windows, en OpenCV.

Chapitre 2 : Présentation du Laboratoire L3i

2.1. Son Histoire

Le laboratoire L3i est reconnu par le Ministère de la Recherche comme Équipe d'Accueil (EA2118) depuis 1997.

Il regroupe désormais une quarantaine d'enseignant-chercheurs (dont 12 habilités à diriger des recherches), essentiellement issus de la communauté informatique (section 27 du CNU) et génie informatique (section 61) de l'Université de La Rochelle. Au premier janvier 2004, 20 doctorants sont inscrits.

Le laboratoire a changé de nom : d' « Informatique et Imagerie Industrielle », il s'est tourné vers « *Informatique, Image, Interaction* ».

Il se positionne donc beaucoup plus sur une image, facteur d'interactions que sur les aspects technologiques associés à une informatique Industrielle. Cette orientation a été motivée par une volonté d'aborder, de la manière la plus cohérente, complète et fondamentale, les différentes facettes de chaque contexte applicatif auquel le laboratoire est confronté. En particulier, l'*ouverture du laboratoire aux autres domaines disciplinaires* (les arts, le littoral, les usages) a montré que la prise en compte des interactions, sous toutes les formes qu'elles prennent, était primordiale.

L'*Image et le Comportement* sont donc au centre de l'activité du Laboratoire.

Une partie de nos activités concerne donc naturellement les aspects analyse/modélisation/synthèse – c'est l'objet de la thématique *Image et séquences d'Images (ISI)*. Pour l'analyse d'Image et pour l'interprétation de comportements complexes, un travail de fusion/classification est nécessaire. La thématique *Données, Formes, Interprétation (DoFIIn)* traite ce problème et le replace dans un cadre plus général, lui permettant d'aborder, en particulier, les données semi-structurées. La thématique *Modèles, Comportements, Architectures (MoCA)* contribue à inscrire les travaux menés dans un contexte opérationnel et un cadre méthodologique. Ceci permet de mieux comprendre, à travers leur modélisation, les systèmes représentés par des images, mais aussi, de produire les logiciels associés à nos résultats en les intégrant dans une architecture pertinente.

Les quatre dernières années ont permis :

- De provoquer des *synergies* entre l'IUT et l'UFR de Sciences afin d'unifier les chercheurs en Informatique de l'Université de La Rochelle. Ceci a permis de restructurer les recherches autour de thématiques dans lesquelles tous les chercheurs ont leur place. Il en résulte que le L3i est désormais le plus gros laboratoire de l'Université de La Rochelle
- de *développer des contacts industriels* permettant de motiver des recherches fondamentales et appliquées.
- de *s'insérer dans les réseaux de recherche*, participer aux réseaux industriels de nos domaines, de préparer l'insertion dans les réseaux européens.
- d'améliorer la *diffusion scientifique*, en qualité et en quantité.

- de trouver une “liberté financière” en *multipliant par quatre les recettes du laboratoire*, tout en limitant les concessions sur la nature des recherches.
- de mener une *politique de recrutement* coordonnée et de limiter les déstabilisations consécutives à des recrutements relativement nombreux, d'accroître le nombre de doctorants (6 par an depuis les 2 dernières années).
- de mettre en place une *politique de valorisation* avec dépôt de brevets et transfert de technologie.

2.2. Sa structuration

Au niveau de la dynamique du laboratoire, afin de ne pas souffrir, comme beaucoup de laboratoires, des barrières induites par des structurations administratives et scientifiques, une réflexion a été menée sur une structuration adaptée à sa taille et correspondant à la dynamique interne du laboratoire. L'activité de la recherche peut ainsi être décrite à l'aide d'une matrice permettant de croiser les thématiques scientifiques et les projets :

- Une thématique correspond à une vision académique d'une problématique scientifique. Elle concerne une communauté de recherche identifiée. Un chercheur appartiendra donc, de manière privilégiée à une thématique. C'est le vecteur dominant. Cependant, les activités de recherches pluri-thématiques constituent un moyen d'organiser les recherches situées aux interfaces entre les thématiques. Elles sont justifiées dès lors qu'un chercheur s'investit dans un domaine disciplinaire et aborde les aspects méthodologiques ou s'il développe une théorie fondamentale, qu'il spécialise dans un contexte disciplinaire particulier.
- Un projet est le support de la transversalité entre thématiques scientifiques. Chaque projet ne se résume pas à un contexte applicatif. On a plutôt considéré des « objets », caractérisés par des contraintes particulières et une classe de problèmes associés, sur lesquels il était nécessaire de capitaliser des connaissances. On discrétise les recherches en fonction des différents problèmes à aborder. Chaque thématique est alors le cadre pertinent pour maîtriser l'ensemble des recherches menées dans le contexte des projets associés. Il est ainsi possible de replacer les travaux en utilisant le corpus des résultats disponibles, de développer de manière la plus générale possible de nouveaux résultats et enfin d'identifier les recherches spécifiques à mener.

Les cellules de la matrice ainsi formée contiennent des Actions qui correspondent à des recherches menées par des chercheurs, spécialement regroupés pour atteindre l'objectif visé. En fonction de cet objectif et de son envergure, une action peut impliquer plusieurs thématiques ou plusieurs projets. Une action est classiquement associée à un contrat avec un partenariat impliquant des industriels et/ou des académiques.

Le financement est alors principalement externe. Une action peut aussi consister en une recherche prospective, définie en interne au laboratoire (après validation par le conseil scientifique). Le financement est alors pris en charge sur fonds propres du laboratoire. De telles actions correspondent à des recherches exploratoires menées dans le laboratoire. Naturellement, des partenaires peuvent être associés. Les recherches prospectives se prolongent naturellement en contrats.

Il apparaît alors clairement que cette structuration induit une dynamique toute particulière puisque chaque action nouvelle est l'occasion, dans le cadre d'un séminaire interne puis, au sein du conseil scientifique, d'identifier les problématiques scientifiques sous-jacentes

(thématiques) et les caractéristiques particulières à prendre en compte (projets). En outre cette organisation permet de provoquer les synergies nécessaires entre chercheurs pour atteindre les résultats visés. Les équipes sont ainsi brassées, de sorte à ce que chacun puisse intervenir, dans son domaine de recherche, dans un contexte différent qui peut susciter de nouvelles recherches et l'enrichir du point de vue des autres chercheurs.

Il serait donc réducteur d'assimiler la notion d'équipe à celle de thématique, et tout aussi faux que d'adopter le point de vue des projets. Chaque chercheur se positionne naturellement dans la matrice en fonction de ses compétences. Thématiques et projets ne sont donc pas des entités administratives, mais scientifiques. Chacune est animée par un responsable qui gère uniquement l'animation et la cohérence des activités scientifiques menées au sein de son entité. La coordination générale et la gestion administrative sont assurées au niveau de la direction du laboratoire.

Ce mode de fonctionnement est possible, d'abord parce que le laboratoire est d'une taille idéale : pas trop important pour imposer une structuration trop lourde, suffisamment riche pour aborder des problèmes complexes à l'aide de compétences multiples. On permet aux chercheurs de se concentrer sur les tâches scientifiques, en les dégageant au mieux des charges administratives. Le dynamisme des chercheurs est aussi un atout prépondérant dans la réussite de ce mode de fonctionnement. Enfin, ce mode de fonctionnement, basé sur l'interaction entre chercheurs, facilite l'ouverture vers des partenaires extérieurs au laboratoire.

2.3. Ses thématiques scientifiques

2.3.1. Image et séquences d'Images (ISI)

- Détection et suivi d'objets sur des séquences d'images ;
- Étude et modélisation de déformations ;
- Restauration et extraction de l'information ;
- Reconstruction, production et suivi de trajectoires ;
- Interaction et planification de séquences de traitement d'images.

Problématique :

La finalité de cette thématique s'inscrit dans la problématique du traitement de l'information, restreint aux composantes « Traitement du Signal, Traitement de l'Image, Reconnaissance des Formes, Modélisation, et Conception d'algorithmes ». La thématique est, par ce fait, en parfaite cohérence avec les deux autres thématiques du laboratoire. De manière conjointe avec celles-ci, les travaux ont pour finalité d'apporter des solutions originales et robustes sur le plan algorithmique : l'information pertinente extraite est combinée dans l'intention :

1. de fabriquer des paramètres sur différents niveaux sémantiques, permettant à des techniques de décision multi-composantes d'évaluer la sortie de la chaîne de traitement.
2. de définir des invariants ou signatures (spatio-temporelles) servant d'indices pour l'indexation d'images ou de documents dans des systèmes d'informations ;
3. de développer une interaction dynamique entre les différents niveaux de la chaîne d'information pour optimiser la chaîne de traitement et d'analyse, cibler les traitements selon les classes d'objets, définir des stratégies dépendantes du contexte et inférer des attributs non accessibles (occultation, image dégradée).

La thématique "image et séquences d'images" s'est d'abord focalisée sur la problématique du traitement des images acquises à partir de systèmes d'imagerie de type industriel. Cette problématique s'est donc déclinée très rapidement sur le mode restauration, extraction d'informations pertinentes et détection selon les spécificités industrielles : conditions d'acquisition difficiles, contraintes d'éclairage, contraintes de temps réel, image de haute définition. Cette thématique s'est par la suite enrichie de réflexions portant sur des domaines d'imageries multi-composantes (couleur : films anciens, images sous-marines, contexte routier, multi-canaux : imagerie satellitaire, multi-modalités : imagerie biomédicale, multi-formes : document (image, graphique, texte)) et s'appuyant sur des séquences temporelles (films anciens, séquences de trafics routiers, séquences de vues d'aquarium, séquences de mouvements dansés, imagerie biomédicale. . .) et/ou sur des contextes de systèmes d'informations documentaires (patrimoine, SIG, . . .). Le triplet restauration, extraction d'information et détection d'objets et se situe résolument au cœur de l'activité thématique dans un contexte de caractérisation et d'interprétation du mouvement, des déformations et des relations spatio-temporelles des objets.

2.3.2. Modélisation, Analyse, Traitement, Recherche d'Informations Complexes et Évolutives (MATRICE)

- Classification robuste,
- Fusion de données à des niveaux d'abstraction et de structuration variables,
- Modélisation pour l'analyse de données spatio-temporelles,
- Modélisation pour une représentation multi-résolution des données spatiales
- Géoréférencées

Problématiques :

Un des objectifs de cette thématique scientifique est de contribuer au développement de systèmes de transformation et de structuration d'informations complexes évolutives. Les données d'entrée de ces systèmes peuvent être de niveau de structuration variables (du faiblement structuré à la base de données), être numériques et/ou symboliques, et surtout concernent des processus évolutifs, pour lequel l'aspect temporel nécessite d'être modélisé (aspect dynamique de la mise à jour des données). Tous les éléments de la chaîne sont intégrés dans ces systèmes, depuis l'analyse des données, jusqu'à leur présentation à un utilisateur émetteur de requête. Les problématiques scientifiques sous-jacentes concernent la classification, la modélisation pour l'analyse mais aussi pour la présentation des données, et l'extraction de caractéristiques, pour le développement de systèmes de recherche d'informations. Modéliser un système d'analyse et de recherche d'informations nécessite une indexation, c'est à dire une représentation du contenu de l'information ainsi que des connaissances du domaine couvert par le corpus, et une correspondance entre la requête et la représentation du document numérique analysé, correspondance permettant un calcul de la pertinence de la réponse. Sur un plan applicatif, les corpus d'analyse sur lesquels les recherches ont été menées concernent essentiellement des images, des séquences d'images ou des données environnementales.

S'appuyant sur des liens étroits avec les thématiques ISI et GéloaMemoLé, les recherches portent sur le développement d'outils d'analyse d'images et d'extraction de primitives, la

recherche de méthodes robustes de classification pour l'indexation et l'interprétation, la fusion d'informations à divers niveaux d'abstraction (données, primitives, décisions) et de structuration (numérique/numérique, numérique/symbolique, symbolique/symbolique). Pour les données semi-structurées, on traite le cas particulier de la modélisation de données spatio-temporelles et de leur interaction avec les données environnementales, de données géo-référencées dans un contexte mobile.

2.3.3. Génie Logiciel - Architecture, Méthodes Modèles, Outils Langages, Évaluation (GéLoAMéMOLÉ)

- Modélisation de comportements d'agents et de leurs interactions dans un contexte de répartition et de mobilité.
- Vérification de leurs propriétés comportementales.
- Définition de modèles et de méthodes pour évaluer et contrôler la qualité de spécifications et du logiciel.
- Conception d'architectures logicielles à composants fiables et réutilisables.
- Modélisation de systèmes d'informations, d'informations complexes et de leur transfert.

Problématique :

Les objectifs de la thématique sont la conception de modèles, de méthodes, de langages et d'outils pour représenter, analyser et mettre en œuvre le comportement des systèmes complexes, mobiles et évolutifs.

Le point de vue du L3i est de systématiquement développer une approche générique, dans chacun de ses secteurs d'activité informatique : modélisation, méthodologie, prototypes. C'est à dire que les modèles, les stratégies, les méthodes et les outils sont conçus dans le souci de permettre leur adaptation et leur réutilisation pour des cas d'étude, des domaines applicatifs et des contextes différents. Un autre point de vue systématiquement adopté est de chercher à garantir la qualité des méthodes, analyses adoptées ainsi que des produits logiciels fournis. Cela passe par un souci de spécification et de documentations précises des domaines, des besoins, des méthodes et des solutions logicielles.

Cette spécification permet de valider et vérifier par le test et le raisonnement formel les applications mises en œuvre dans les prototypes. Elle répond également à la volonté mentionnée plus haut de travailler sur des solutions génériques réutilisables.

Les activités de recherche sont organisées en trois classes de préoccupations complémentaires :

Modélisation et conception : Intégration de modèles, modèles de comportement, définition de composants, validation de composants, automatisation de production de code ou de modèles intermédiaires, modélisation à base de systèmes multi agents, modèles de données pour les systèmes de visualisation d'informations spatiales embarquées, modélisation des systèmes d'informations géoréférencées.

Démarche de conception : Travaux sur la cohérence métriques des modèles produits et sur la méthodologie de production, critères, métriques et stratégie qualité.

Production d'outils théoriques et cadre méthodologique : treillis de Galois et ordres partiels, systèmes implicatifs.

Ces préoccupations définissent les activités de chercheurs du laboratoire. Chaque préoccupation se décline en différentes problématiques.

Chapitre 3 : Indexation et recherche d'images [1]

3.1. Introduction

Aujourd'hui la recherche d'images ressemblantes dans des bases de données ne se limite plus à l'indexation textuelle car cette approche est assez contraignante ; En effet, la description textuelle d'une image demande une intervention humaine ; et celle-ci a deux limites principales : le temps et le jugement dans parler des limitations liés à la langue utilisée pour l'indexation.

D'une part, présenter le contenu d'une image demande beaucoup de temps. Pour des ensembles d'une certaine d'image cela encore imaginable mais actuellement les volumes étudiés sont plutôt de l'ordre des milliers voire de la centaine de milliers d'images. Une intervention humaine pour décrire de tels ensembles semble donc impossible.

D'autre part, la présentation qui est faite d'une image peut varier énormément suivant la personne. Les différences peuvent porter sur :

- La langue utilisée : anglais, français, allemand,...
- Le vocabulaire utilisé : une Porsche sera décrite par « Porsche 911 noire » par un connaisseur mais par « voiture se sport noire » dans la plupart des cas.

Par conséquent, ne pouvant pas se baser entièrement sur une description textuelle, la recherche d'image tente aujourd'hui d'extraire les informations directement des images. Ce type de travaux se regroupent sous le terme : indexation d'images par le contenu (ou CBIR en anglais pour **C**ontent **B**ased **I**mage **R**etrieval). Dans ce domaine les images sont décrites grâce à leur propre contenu. Elles sont analysées afin d'en extraire les éléments permettant de les caractériser le plus fidèlement possible.

Ce chapitre se propose de présenter, tout d'abord, les différentes questions qui se posent en indexations d'images. Ils amènent ensuite à trouver un système de représentation d'images adapté au problème. Nous exposerons alors un rapide aperçu des différentes techniques les plus employées en recherche d'images par le contenu ; tout d'abord dans un schéma où l'image est décrite dans sa globalité puis en prenant en compte l'organisation spatiale des objets au sein de l'image. Enfin nous présentons les diverses méthodes utilisées pour quantifier la qualité de tels systèmes de recherche.

3.2. Représenter pour indexer et rechercher

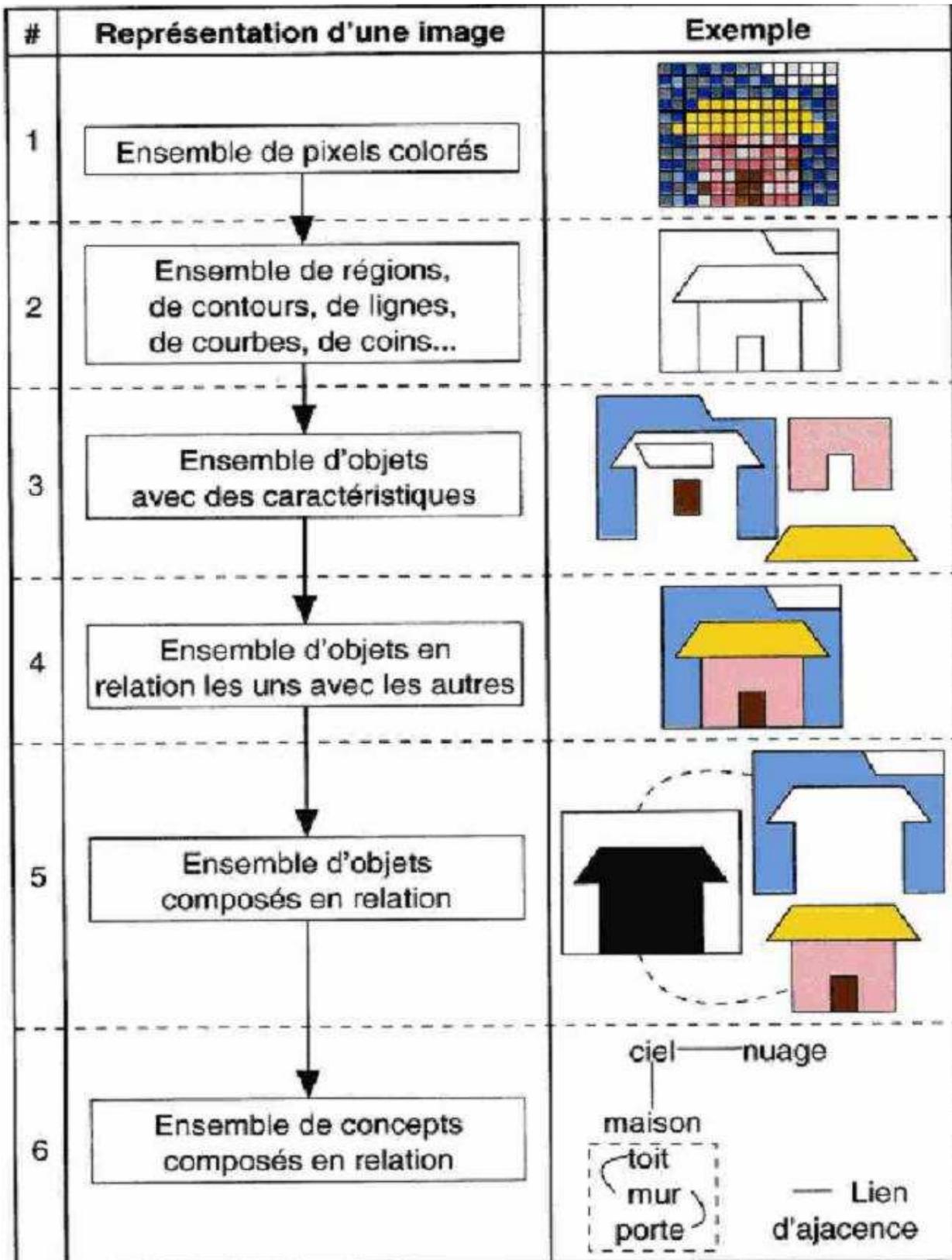
3.2.1. Modes de recherche possibles :

Globalement l'indexation d'images par le contenu a pour but de retrouver des images d'un contenu bien précis. Tout l'enjeu est que l'information recherchée peut être de nature très diverse. La liste ci-dessous donne un panel des problèmes qui peuvent être posés :

- Description simple : l'utilisateur désire retrouver des images à partir d'un critère simple et global à l'image. Pour retrouver des images de plages, la demande pourra être formulée en terme de couleurs : « Trouver les images contenant 50% de bleu, 30% de vert et 20% de jaune ».
- Description composée simple : les images sont recherchées suivant de multiples critères combinés. L'obtention d'images de forêt pourra être réalisée par : « Trouver les images à dominante verte et contenant une texture semblable à celle d'un arbre. Le critère de couleur étant plus important que l'autre ».
- Recherche par esquisse : la recherche s'effectue de la même manière que dans le cas précédant mais la référence est un dessin réalisé par l'utilisateur et non une image.
- Description localisée : l'utilisateur spécifie la position des critères qu'il propose. La demande « Trouver les images contenant du bleu en haut et du jaune en bas » pourra être défini pour rechercher des images de plages.
- Recherche par l'exemple : dans ce cas une image est utilisée comme référence pour effectuer la recherche. Ainsi, les demandes effectuées seront fonction des caractéristiques de l'image de référence voire même fonction d'une partie de celle-ci ; Par exemple, il est possible de rechercher les images contenant des arbres en sélectionnant un arbre dans une image et en demandant « les images de composition colorimétrique et de texture semblables à celle de cet arbre ».
- Description d'un objet : un problème plus délicat est la recherche d'objets au sein des images. Dans ce cas, l'utilisateur peut vouloir obtenir « les voitures rouges positionnées au centre de l'image ». Il spécifie alors l'objet qu'il désire en le sélectionnant dans une image ou en l'esquissant et en définissant des critères sur cet objet et /ou sur son positionnement : « forme de voiture », « couleur rouge » et « au centre de l'image » pour l'exemple précédent ce type de demandes est appelé requête partielle.
- Description d'un ensemble d'objets : l'évolution de la requête précédente est de vouloir obtenir un ensemble d'objets organisés d'une certaine façon. Par exemple : « Trouver les images d'une route sur laquelle une voiture rouge suit une voiture bleue ».

Il apparaît que les problèmes posés sont très variés et que les solutions apportées peuvent être très diverses suivant les contextes, en particulier le type de bases (généraliste ou spécialisée) et le type d'utilisateurs (expert du domaine ou non).

3.3. Représentation d'une image :



Niveaux d'abstraction auxquels une image peut être vue

Les premières approches développées en indexation décrivaient les images dans leur ensemble. Des statistiques simples sont alors calculées pour représenter les images. Généralement, les informations prises en compte sont la couleur, la texture et/ou la forme. Malheureusement, ces techniques décrivent les images globalement. Il paraît difficile avec ce type d'approche d'obtenir une description plus fine de l'image et notamment de rechercher des objets.

Depuis quelques années, des approches dites spatiales essaient de fournir une description plus précise en considérant l'image comme composée d'un ensemble d'objets (de régions). La représentation de l'image est alors portée par l'ensemble des descriptions des éléments la composant mais aussi par les relations existantes entre eux. La conceptualisation de l'image a ainsi gagné un niveau.

Nous détaillons maintenant un peu plus l'approche globale puis l'approches spatiale.

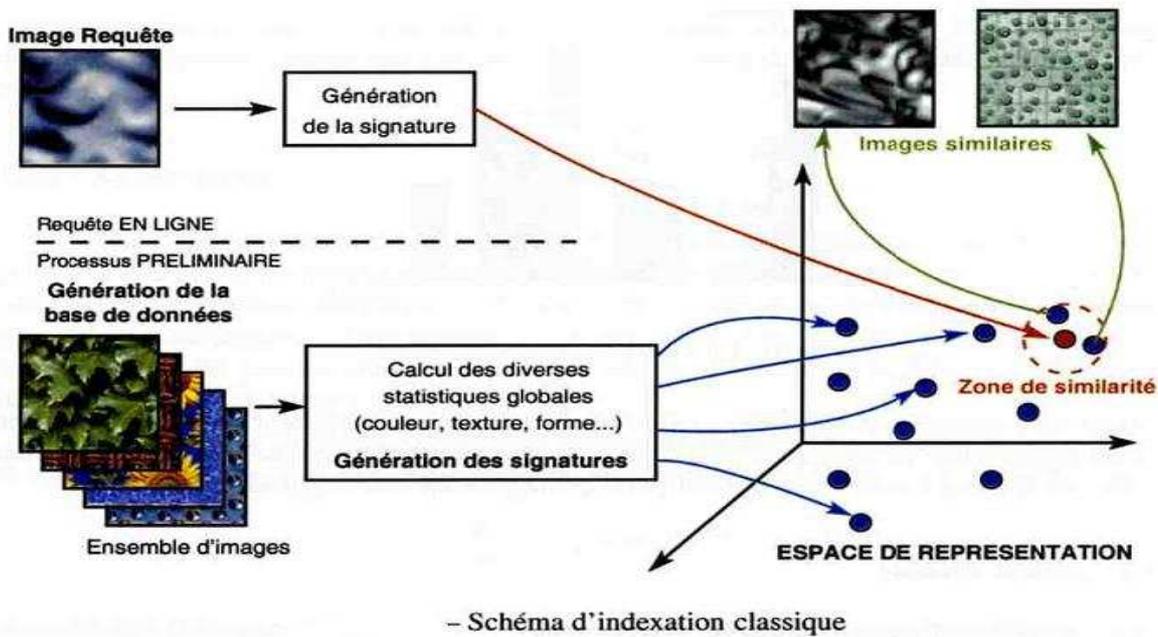
1) Approche globale

Cette approche considère l'image dans son ensemble et la caractérise en utilisant des statistiques calculées sur l'image entière.

Les différentes approches globales portent généralement sur trois critères : la couleur, la texture et la forme. Pour chacun d'eux, un grand nombre d'attributs ont été développées. Ils peuvent ensuite être combinés pour obtenir un descripteur complet et robuste. Cet ensemble de statistiques est généralement appelé signature de l'image.

Ensuite dans le cadre d'une recherche par l'exemple, la démarche est généralement la suivante :

- Phase préliminaire : calculer les descripteurs de chaque image de la base de données ;
- Phase en ligne : calculer les descripteurs de l'image requête ;
- Phase de recherche : rechercher les images proches de l'image requête dans l'espaces du (des) descripteur(s) utilisé(s).



Principaux types de statistiques :

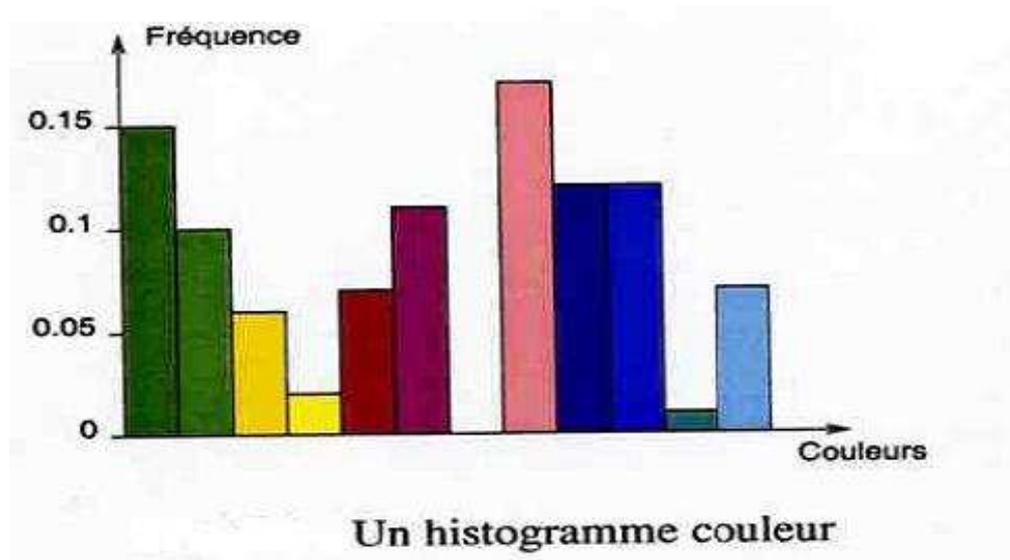
A. Aspect couleur :

La couleur est sûrement le critère le plus important psychovisuellement parlant dans la vision d'une image. Le premier regard porté sur une image d'appréhender très rapidement sa couleur dominante et ses différentes teintes.

La distribution des couleurs au sein d'une image est très souvent représentée sous la forme d'un histogramme couleur. Celui-ci associe à chaque couleur sa fréquence d'apparition. Ainsi, pour une image définie par $I : x \in \Omega \longrightarrow I(x) \in \Theta$ avec Ω le domaine de l'image et Θ celui des couleurs ; la fonction histogramme de l'image I peut s'écrire $H_I : c \in \zeta_I \longrightarrow \mathfrak{R}$ où $\zeta_I = \{ I(x) / x \in \Omega \}$ est l'ensemble des couleurs de l'image (ce qui entraîne que $nb_{\zeta_I} = \text{Card}(\zeta_I)$ est le nombre de couleurs présentes au sein de l'image I) les valeurs de l'histogramme sont alors :

$$H_I(c) = \frac{\text{Card}(\{x / x \in \Omega, I(x) = c\})}{\text{Card}(\Omega)}, \forall c \in \zeta_I. \quad \sum_{c \in \zeta_I} H_I(c) = 1$$

Il est noté ici qu'avec cette définition l'histogramme obtenu est normalisé,



B. Aspect forme :

Une fois l'analyse colorimétrique effectuée, le cerveau a décomposé la scène en plusieurs régions. La forme peut, par la suite, être utile pour les décrire. Elle est d'ailleurs très utilisée en vision industrielle et notamment pour du tri de pièces ou de la vérification de placement de composants électroniques. En prenant cette image :



Figure 4.

Le poivron et le concombre sont colorimétriquement très proches alors que leur forme est très différente ; c'est le critère qui doit être utilisé pour les distinguer.

La forme d'une région est décrite sous la forme d'une image I de taille $[l, h]$. Cette image définit alors la boîte englobante de l'objet. Les points de l'image I sont tous les couples (x, y) où $x \in [0, l-1]$ et $y \in [0, h-1]$.

Le masque définissant la forme est alors la fonction :

$f(x,y) : (x,y) \in I \rightarrow B$ où $f(x,y) = \text{vrai}$ si le point (x,y) fait partie de la forme.

De même la frontière de la région est écrite par la définition de la fonction :

$g(x,y) : (x,y) \in I \rightarrow \mathbb{B}$ où $f(x,y) = \text{vrai}$ si le point (x,y) fait partie de la frontière.

La reconnaissance des formes a débuté en appliquant des méthodes issues de la caractérisation de variables aléatoires et en particulier les moments. Les moments $p+q$ sont définis par :

$$m_{pq}(I) = \sum_{(x,y) \in \Omega} x^p y^q I(x,y), \quad \forall (p,q) \in \mathbb{N}^2$$

Où I est l'image au sein de laquelle la forme à étudier a été isolée. Ainsi, les moments d'ordre 1 normalisés par m_{00} , $\frac{m_{01}}{m_{00}} = \bar{y}$ et $\frac{m_{10}}{m_{00}} = \bar{x}$, correspondent à la position du centre de gravité de

la forme. Les moments centrés peuvent être définis en utilisant \bar{x} et \bar{y} :

$$\mu_{pq}(I) = \sum_{(x,y) \in \Omega} (x - \bar{x})^p (y - \bar{y})^q I(x,y), \quad \forall (p,q) \in \mathbb{N}^2$$

Les différentes valeurs de μ_{pq} permettent de caractériser une forme ; μ_{01} mesure par exemple l'allongement suivant y . Les formes de deux objets sont alors comparables simplement grâce à leurs moments respectifs.

Hu propose en 1962 un ensemble de combinaisons des moments centrés définissant 7 nouveaux moments invariants par translation, rotation et changement d'échelles.

De nombreuses autres techniques ont aussi été mises en place pour caractériser une forme est la différentiel d'une autre : descripteurs de Fourier, moments de Zernike, représentation CSS.

C. Aspect texture :

La texture est un concept très lié à la perception mais assez difficile à caractériser du point de vue de la vision. Cette caractéristique est plutôt associée au toucher ou au goût. La notion de rugosité est ainsi difficile à analyser sur une image 2D. Par contre, le contraste paraît un peu plus naturel, la notion de texture peut sommairement être assimilée à la description du (des) motif(s) formé(s) par les différentes couleurs : uniformes, alterné une ligne sur deux...

Dans ce cadre, HARALIK propose d'utiliser la matrice de cooccurrences [HARALI73]. Celle-ci contient les informations des voisinages entre les différentes couleurs composant l'image.

Plus précisément, la fonction de cooccurrences au sein d'une image I est définie telle que

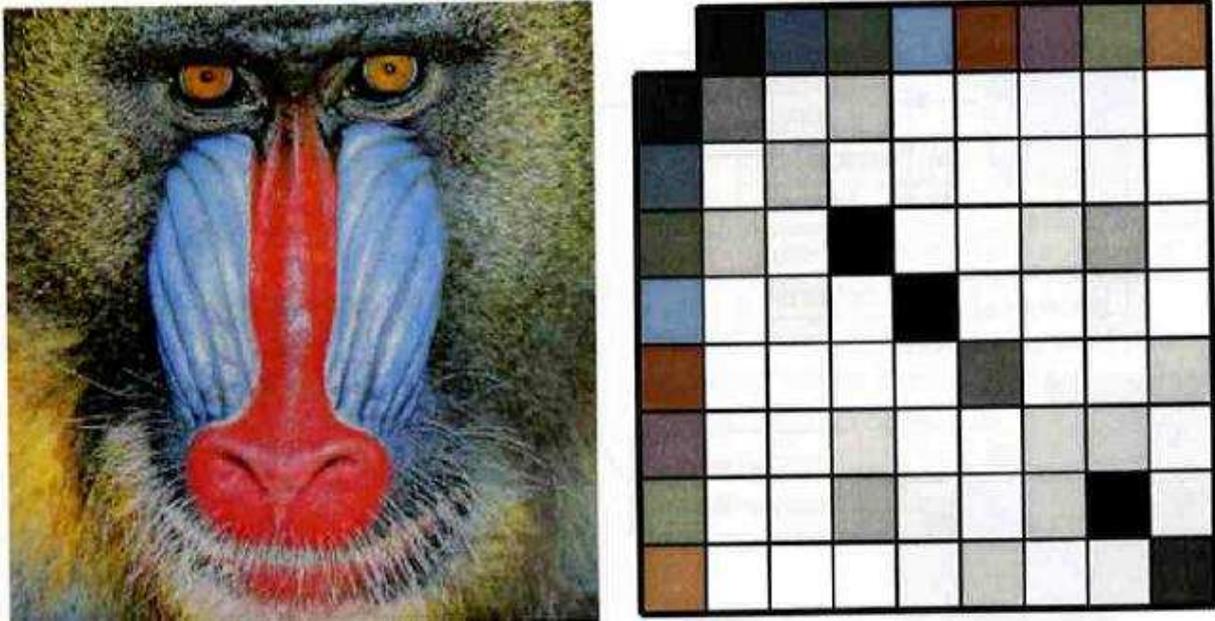
$MC_I : \zeta_I^2 \rightarrow \mathbb{R}$. de couleur $(c_1 \text{ et } c_2)$ est associée la probabilité qu'elles soient voisines. Il faut donc définir une fonction de voisinage : qui définit si $(p_1, p_2) \in \Omega^2$ sont voisins suivant le cas $V : \Omega^2 \rightarrow \mathbb{B}$ 4-connexité, 8-connexité ou autre (pour une direction et une longueur donnée par exemple).

Ainsi, $\forall (c_1, c_2) \in \zeta_I^2 :$

$$MC_I(c_1, c_2) = \frac{\text{Card}(\{(p_1, p_2) / (p_1, p_2) \in \Omega^2, I(p_1) = c_1, I(p_2) = c_2, V(p_1, p_2) = \text{vrai}\})}{\text{Card}(\{(p_1, p_2) / (p_1, p_2) \in \Omega^2, V(p_1, p_2) = \text{vrai}\})}$$

La matrice de cooccurrences est égale à la matrice carrée de taille $nbc_I = \text{Card}(\zeta_I)$:

$$[MC_I] = [MC_I(c_1, c_2)]$$



Matrice de cooccurrences

Figure 5.

A partir de celle-ci, HARALICK définit 14 mesures pour caractériser une texture telles que le contraste, la corrélation, la variance...

D'autres statistiques ont été mises en place à partir de cette même matrice de cooccurrences mais aussi à partir des coefficients d'ondelettes ou GABOR ou encore avec la définition de matrices de longueurs de plages.

2) Approches spatiales

L'idée pour cette approche est de considérer l'image comme un ensemble d'objets et non plus comme une entité unique. Une fois l'image segmentée en plusieurs régions, ces dernières peuvent être caractérisées des mêmes manières que les images dans la partie précédente. De cette façon, des détails plus fins peuvent être retrouvés au sein des images car les statistiques sont localisées ce qui induit une diminution du problème de masquage d'une donnée par une autre (la composition d'un objet ne sera plus noyée dans celle de l'image mais sera isolée). De plus, il est possible de rechercher des arrangements spécifiques d'objets en utilisant l'information liée à l'organisation des régions dans l'image.

Ensuite dans le cadre d'une recherche par l'exemple, la démarche est généralement la suivante :

- Phase préliminaire : segmentation des images de la base de données et calcul des descripteurs pour chaque région pour obtenir les représentations des images ;
- Phase en ligne : segmentation de l'image requête et calcul des descripteurs des régions pour obtenir la représentation de l'image. Les régions à utiliser pour la recherche peuvent aussi être sélectionnées ;

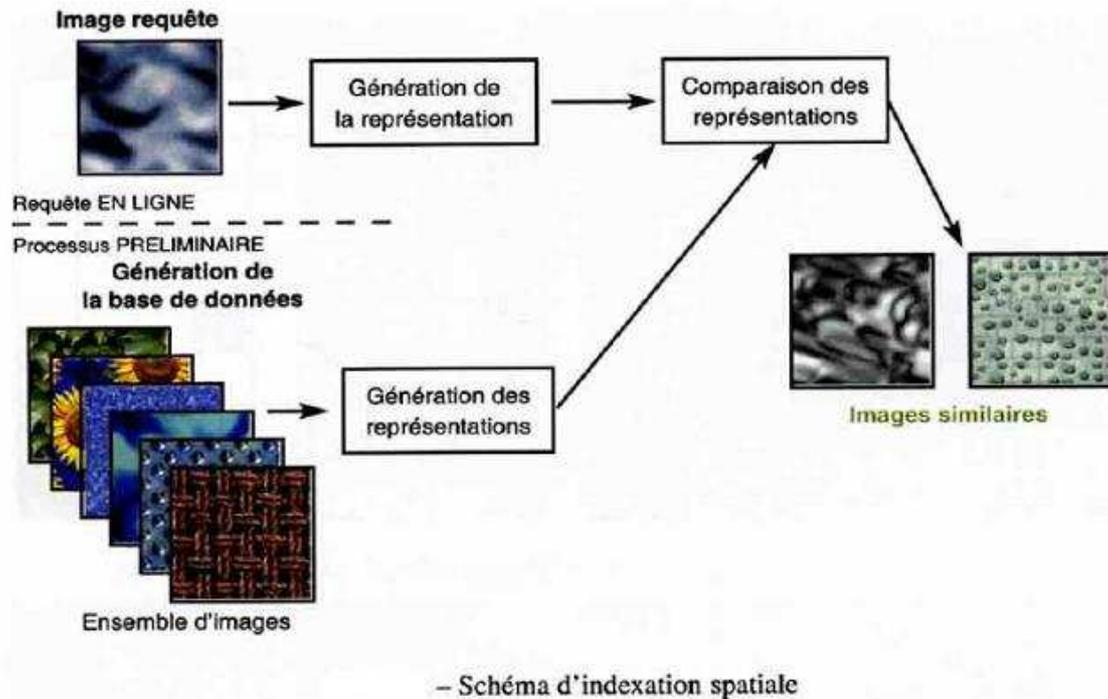


Figure 6.

- Phase de recherche : rechercher les images proches de l'image requête en comparant les descriptions.

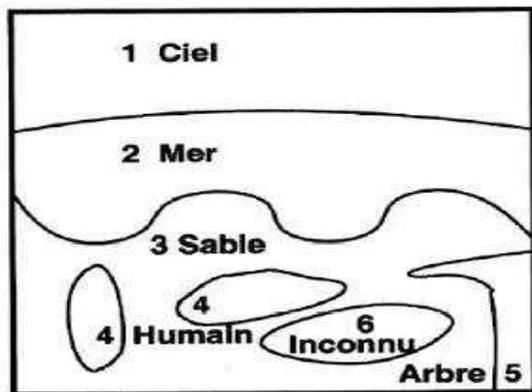
La problématique de l'indexation dite spatiale prend naissance dans la fin des années 1980 et se développe rapidement à la fin du siècle dernier. Dans ce type de méthodes nous ne considérons pas celle qui concernent les requêtes partielle, c'est-à-dire celles qui n'effectuent qu'une recherche de régions. En effet, dans le cadre des requêtes partielle, une fois les images segmentées et les régions décrites, le système de recherche est exactement le même que celui de l'approche globale mais les éléments recherchés ne sont pas les images entières mais certaines régions particulières. Nous donc les systèmes d'indexation spatiale comme ceux prenant en compte l'organisation des différents objets au sein des images pour effectuer le recherche.

Quelques méthodes de représentation d'images :

- Méthode fondée sur les 2-D strings. Les 2-D strings apportent une représentation efficace du contenu d'une image et permettent de réduire la complexité de la comparaison. Chaque objet est identifié avant le stockage, un identifiant unique devant leur être affecté. Les positions relatives entre les objets sont alors représentées par deux chaînes de caractères

mono- dimensionnelles. L'idée est de projeter la position des objets (i.e. leurs centres de gravité) sur les deux axes de l'image et de les prendre dans le même ordre que leurs projections. Le problème de recherche est alors transformé en une comparaison de chaînes des caractères.

- Méthode utilise des caractéristiques autres que spatiale telles que l'aire ou le périmètre des régions.
- Méthode utilise le graphe d'adjacences de régions. Les régions ne sont plus nécessairement identifiées et peuvent être décrites par les diverses statistiques utilisées en indexations globale ; les arrangements spatiaux décrits peuvent être plus complexes qu'un simple ordonnancement vertical et horizontal.



À gauche l'image d'exemple et dans le tableau ci-dessus des exemples de 2-D strings qui peuvent être utilisés. Les régions de l'image ont été au préalable segmentées et identifiées.

– Exemples de 2-D strings

Figure 7.

3.4. Comparaison des descripteurs :

SWIN et BALLARD ont les premiers proposés une méthode d'intersections d'histogrammes permettant de comparer des images par leur distribution colorimétrique. Considérant deux images I_1 et I_2 et leurs histogrammes respectifs H_1 et H_2 définis sur le même domaine couleur ζ , la distance proposée est :

$$D_{SwainBallard}(H_1, H_2) = \frac{\sum_{c \in \zeta} \min(H_1(c), H_2(c))}{\sum_{c \in \zeta} H_1(c)}$$

Cette mesure revient à trouver l'aire commune aux deux distributions.

Une autre distance utilisée pour comparer deux histogrammes : c'est la distance de MINKOSKI :

$$D_{MINK}^p(H_1, H_2) = \left(\sum_{c \in \zeta} |H_1(c) - H_2(c)|^p \right)^{\frac{1}{p}}$$

Trois formes de cette distance sont très répandues pour la comparaison d'histogrammes couleur :

- L^1 : $D_{\text{MINK}}^1(H_1, H_2) = L_1(H_1, H_2) = \sum_{c \in \zeta} |H_1(c) - H_2(c)|$
- L^2 : $D_{\text{MINK}}^2(H_1, H_2) = L_2(H_1, H_2) = \sqrt{\sum_{c \in \zeta} |H_1(c) - H_2(c)|^2}$
- L^∞ : $D_{\text{MINK}}^\infty(H_1, H_2) = L_\infty(H_1, H_2) = \max_{c \in \zeta} |H_1(c) - H_2(c)|$

La distance d'ordre 1 correspond à la distance L_1 aussi nommé distance de Manhattan, celle de l'ordre 2 correspond à la distance L_2 ou euclidienne et celle d'ordre infini correspond a la distance L_∞ ou de CHEBYCHEV.

- Distance quadratique : $D_{\text{Quadra}}(H_1, H_2) = (h_1 - h_2)^T A (h_1 - h_2)$

A partir de la distance quadratique, de nouvelles distances ont été définies et en particulier celle de Mahalanobis et de Bhattacharyya.

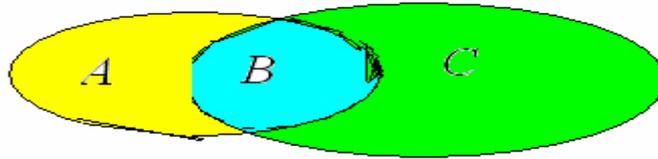
Ces distances sont aussi utilisées pour la comparaison des vecteurs comme étant les descripteurs extraient sur les images de la base de données.

3.5. Evaluer l'indexation :

La performance d'un système d'indexation est très subjective. Généralement, un système peut être comme efficace s'il retourne les informations désirées. La quantification de cette notion est très difficile car le jugement en lui-même est subjectif. Les images retrouvées sont-elles réellement pertinentes ? La réponse sera différente d'une personne à l'autre ; il faudrait donc pour évaluer correctement un système que de nombreuses personnes effectuent de nombreuses requêtes. Le terme « nombreux » apparaissant au moins deux fois de trop sans la phrase précédente, un tel système d'analyse paraît irréaliste à grande échelle. En considérant une base de 1000 images, la performance réelle du système pourrait être évaluée de manière correcte si 100 personnes effectuaient chacune 1000 recherches et qu'elles déterminaient la pertinence des résultats obtenus selon leurs propres critères.

Les méthodes les plus répandues pour évaluer la recherche d'images sont les notions de précision et de rappel. La précision correspond au pourcentage d'images pertinentes retrouvées su sein de l'ensemble des images retournées par le système ; le rappel donne le pourcentage d'images pertinentes retournées par rapport au nombre total d'images

pertinentes dans la base ; En utilisant les notations de la figure, la précision et le rappel sont définis par :



A : images pertinentes non retrouvées.

B : images pertinentes retrouvées

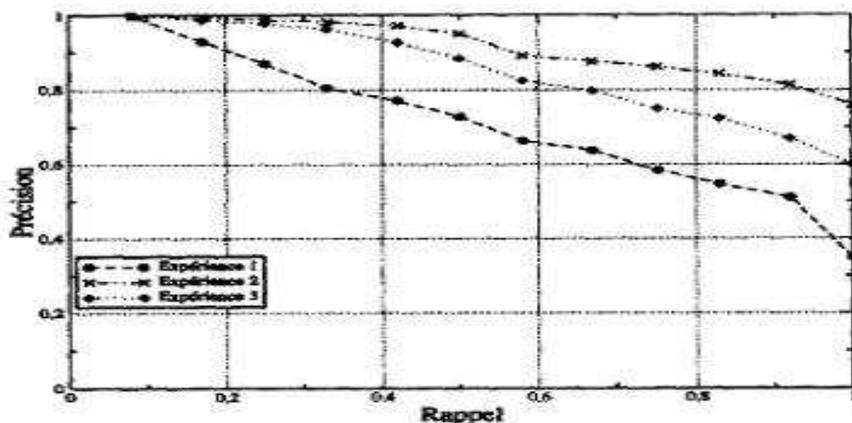
C : images non pertinentes retrouvées.

A+C : images pertinentes.

B+C : images retrouvées.

Figure 8.

Il est possible de tracer le graphe Précision/Rappel. Ce type de courbes correspond à la performance du système pour une requête donnée en fonction du pourcentage d'images pertinentes retrouvées. Un système parfait aura une courbe constante de valeur 1. Plus les courbes tendent vers cet extrême plus les méthodes sont performantes.



– Exemple de courbes Précision/Rappel

Figure 9.

Chapitre 4 : Développement

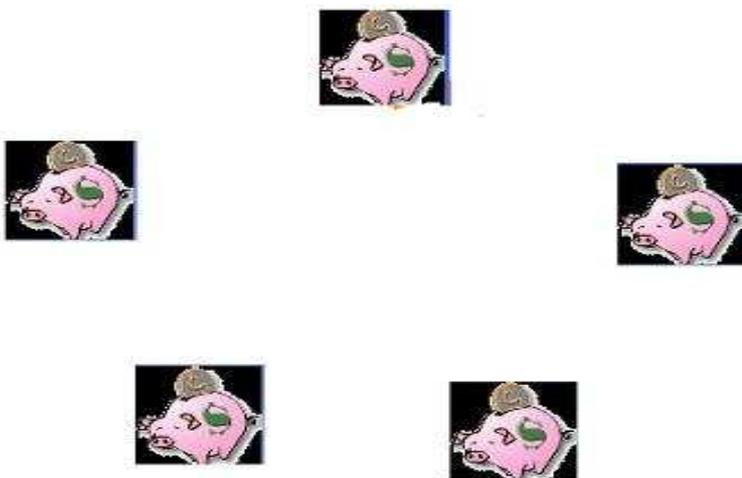
Le cœur de tout système de reconnaissance d'objets est le modèle de description de forme utilisé. De ses propriétés (discrimination, résistance au bruit, invariance aux rotations etc.) vont dépendre les performances du système. Une partition est généralement convenue entre les modèles opérant au niveau des contours et ceux intégrant les régions. Lorsque les objets sont complexes i.e. constitués de régions disjointes (symboles, logos etc.) les techniques fondées exclusivement sur les contours (descripteur de Fourier, sondes circulaires, code de Chain) sont insuffisantes. Dans cette partie nous proposons deux nouvelles mesures captant la topologie d'un objet. Dans le système de coordonnées de l'image tout objet est défini par un ensemble de points d dimensionnel - ayant une topologie spécifique._

Nous adoptons l'approche spatiale puisque, nous semble-t-il, qu'elle est la meilleure représentation du contenu de l'image. (Bien détaillé dans le chapitre précédent).

Les données sur lesquelles seront testés les algorithmes sont des lettrines, caractéristiques graphiques très fréquemment utilisées sur les documents anciens.

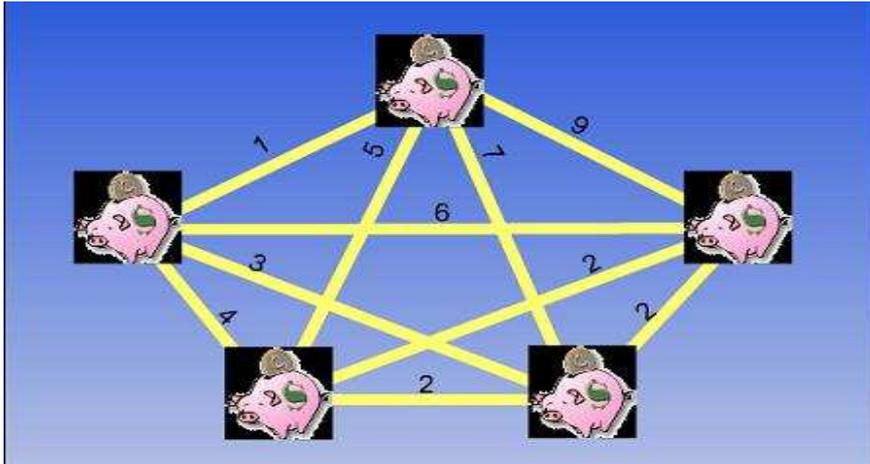
4.1. Première Méthode : la longueur de l'Arbre de Recouvrement Minimale (ARM). En Anglais: Le **Minimum Spanning Tree** (MST).[2]

4.1.1. Définition du MST :



On désire relier toutes les succursales d'une banque et c'est à un coût minimum.

© 2005 Université de La Rochelle



On connaît tous les coûts de chaque connexion entre deux agences.

Alors le MST c'est comment relier toutes les succursales à un coût minimum

4.1.2. Calcule du MST

En connectant tous les points sous la contrainte de définir un arbre de longueur minimale (Minimum Spanning Tree, MST) on capte la topologie de l'objet cible car pour un ensemble de points donnés cette longueur est unique.

Dans la littérature il existe deux algorithmes pour calculer le MST, le premier c'est l'algorithme de KRUSHKAL et le deuxième c'est l'algorithme de PRIM.

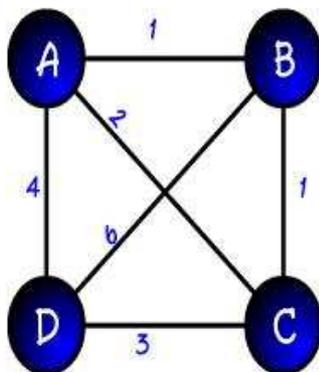
Pour des raisons de facilité d'implémentation, nous avons choisi d'utiliser l'algorithme de PRIM. En effet, ce dernier travaille sur deux ensembles, chacun composé des sommets du Graphe, ensuite, il suffira de faire des comparaisons sur les valeurs de tous les liens existant entre ces deux sous-ensembles. L'algorithme de KRUSHKAL, nécessite la mise en place d'un algorithme permettant de détecter les cycles au sein d'un arbre.

Exemple de calcul de la longueur de MST selon PRIM :

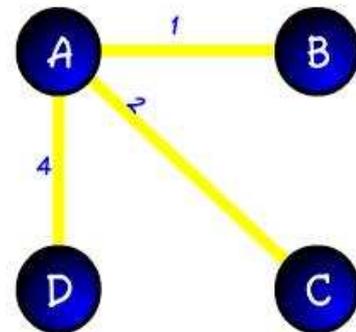
Etape 0.

Etape 1.

PRIM
 On divise le graphe en deux parties, S et X-S.
 On a N = 4 sommets.
 On choisit le sommet A comme sommet initial.
 $|S|=1$
 On incorpore les sommets dans S jusqu'à ce que $|S|=N$.



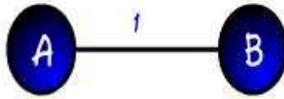
PRIM
 Le cocycle est constitué des arêtes AB, AC, AD.
 Celle de plus petite valuation est AB.
 On incorpore donc B dans S.
 $|S|=2$



Etape 3.

Etape 4.

PRIM



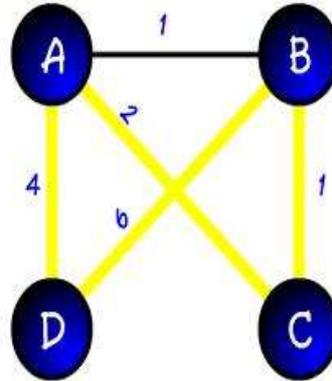
Le cocycle est constitué des arêtes AB, AC, AD.

Celle de plus petite valuation est AB.

On incorpore donc B dans S. |S|=2



PRIM



Etape 2

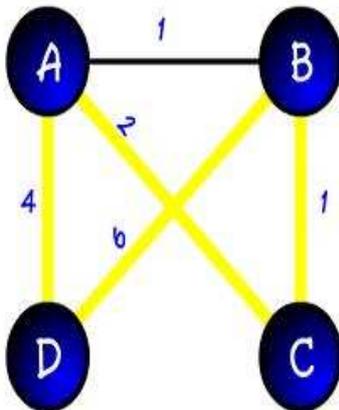
Le cocycle est constitué des arêtes AC, BD, AD, BC.

Celle de plus petite valuation est BC.

On incorpore donc C dans S. |S|=3

Etape 5.

PRIM



Etape 2

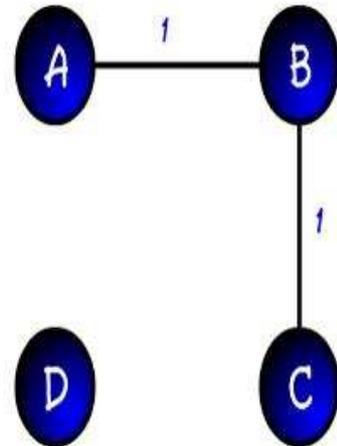
Le cocycle est constitué des arêtes AC, BD, AD, BC.

Celle de plus petite valuation est BC.

On incorpore donc C dans S. |S|=3

Etape 6.

PRIM



Etape 2

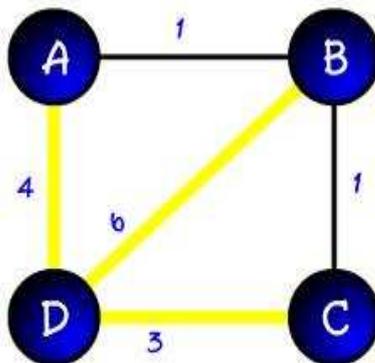
Le cocycle est constitué des arêtes AC, BD, AD, BC.

Celle de plus petite valuation est BC.

On incorpore donc C dans S. |S|=3

Etape 7.

PRIM



Etape 3

Le cocycle est constitué des arêtes BD, AD, CD.

Celle de plus petite valuation est CD.

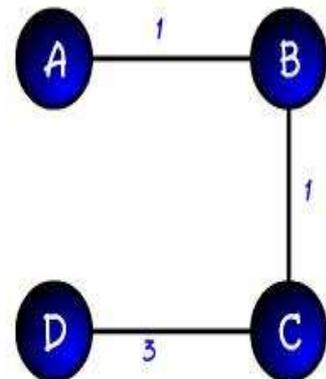
On incorpore donc D dans S. |S|=4

- L'algorithme est fini -

Le coût de l'arbre est : $1+1+3=5$.

Etape 8.

PRIM



Etape 3

Le cocycle est constitué des arêtes BD, AD, CD.

Celle de plus petite valuation est CD.

On incorpore donc D dans S. |S|=4

- L'algorithme est fini -

Le coût de l'arbre est : $1+1+3=5$.

4.1.3. Algorithme pour extraire les signatures en utilisant le MST.

Etapes de l'algorithme :

Entrées : image segmentée

Sorties : la longueur du MST.

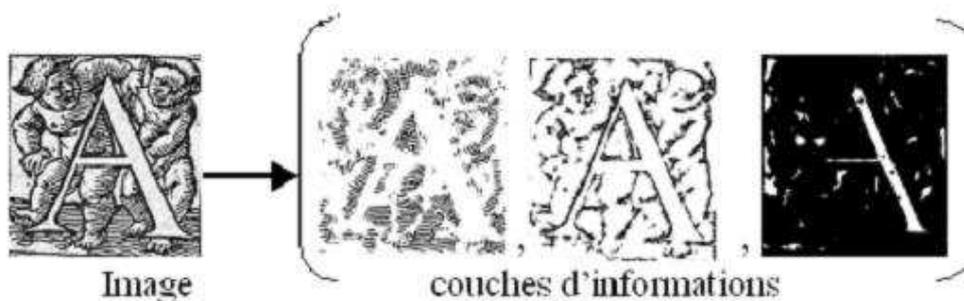
On utilise les graphes pour décrire la composition et la topologie de l'image.

- 1) Segmenter l'image
- 2) Calculer les centres de gravité de chaque région.
- 3) Calcule des distances entre ces centres de gravité.
- 4) Construire le graphe dont les sommets sont les centres de gravité et les arêtes sont les distances calculées entre ces centres.
- 5) Appliquer l'algorithme de PRIM pour calculer la longueur du MST.

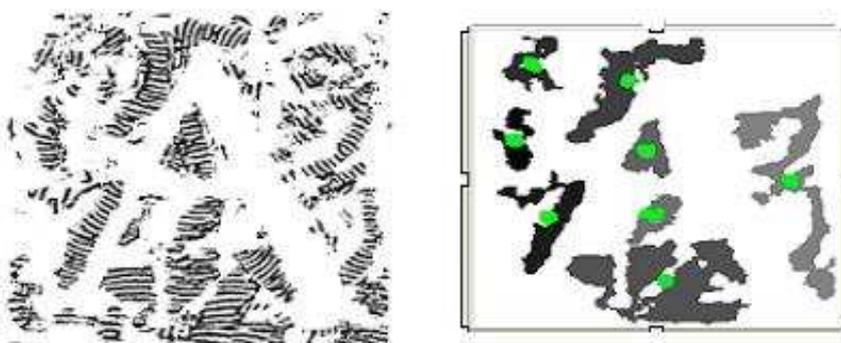
4.1.4. Application sur les lettrines :

- 1) Segmenter l'image :

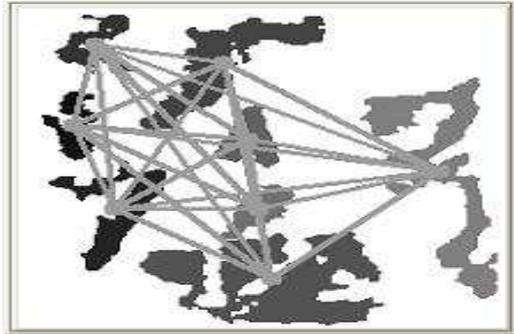
Cette étape consiste à segmenter la lettrine en différentes couches d'informations. Cette segmentation qui est basée sur la technique « first top-down » nous donne plusieurs couches d'informations de l'image originale. Par exemple : zones texturées, zones homogènes, les contours.



- 2) Alors pour chaque couche d'information on calcule la longueur du MST
 - a) Calculer les centres de gravité des régions.

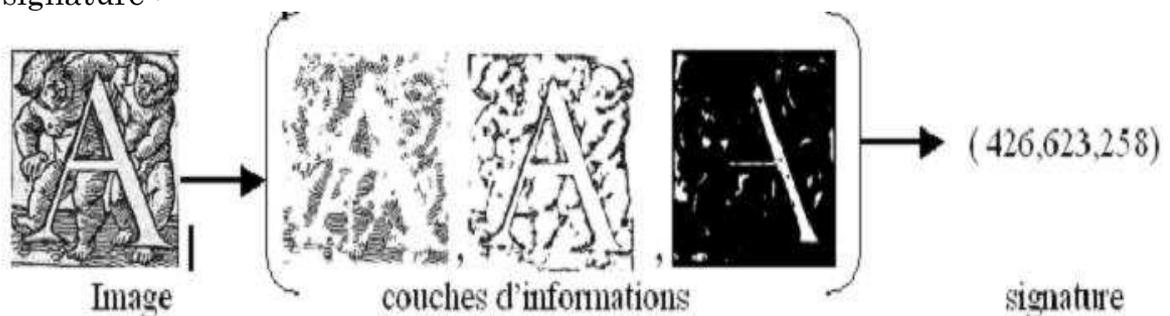


- b) Calculer les distances entre ces centres de gravité
- c) Construire le graphe



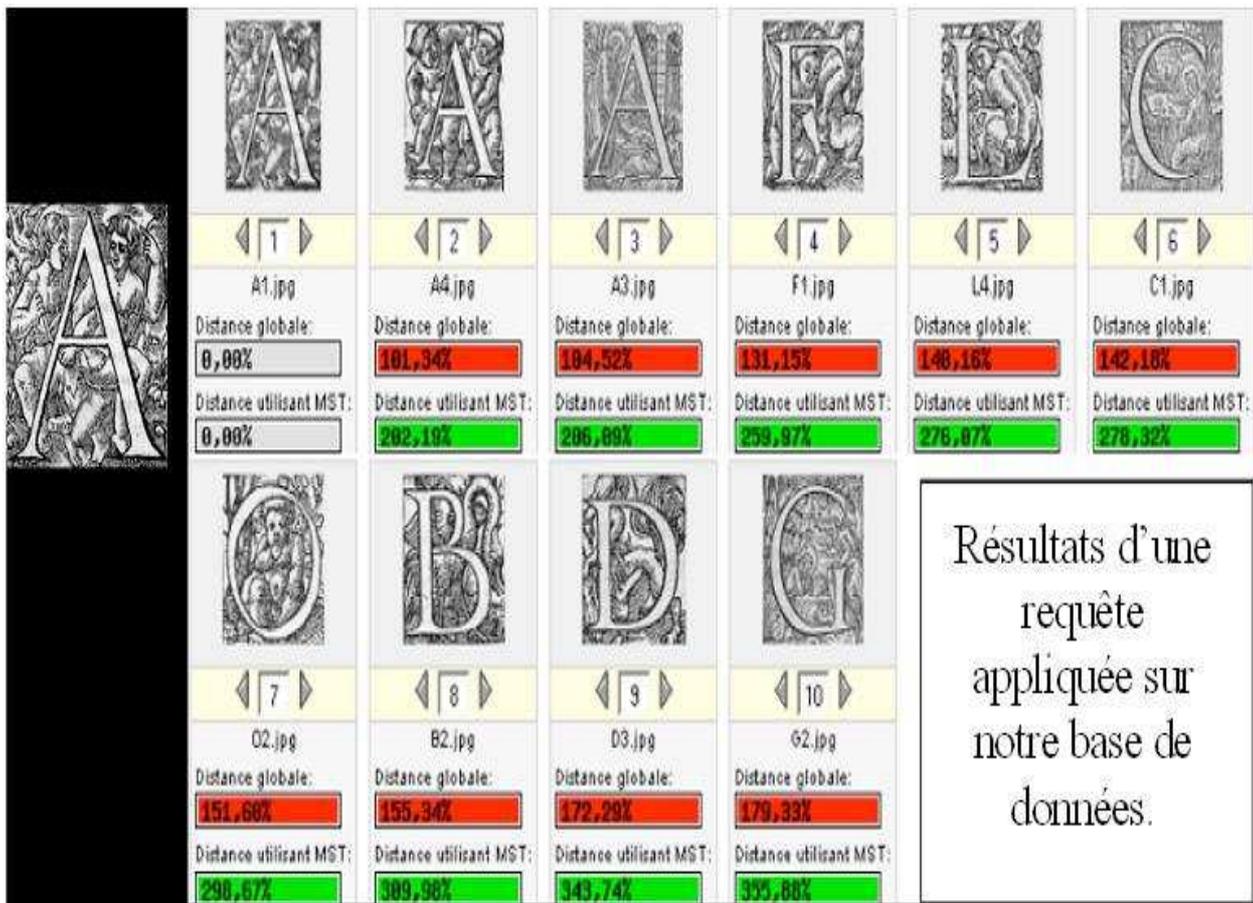
- d) Calculer de la longueur du MST qui est égale à 426.

De même on calcule le MST pour chaque couche d'information. Alors pour chaque image on aura une signature :



4.1.5. Teste :

Cette stratégie de descriptions des images graphique est expérimentée sur une grande base des lettrines en utilisant l'interface graphique de Madonne.



4.1.6. Evaluation :

4.1.7. Problèmes rencontrés

Cette stratégie est évaluée en termes de critère Rappel/Précision. Les premiers résultats ont montré une grande qualité dans la recherche des images par le contenu.

:

Le calcul du MST est alors une fonction dont l'entrée est une image segmentée et la sortie est une valeur : la longueur du MST :

F : base de données $\rightarrow \mathbb{R}$
 Image \rightarrow VAL : longueur du MST.

Cette fonction n'est pas bijective, car une valeur peut être associée à deux images différentes.

Exemple :

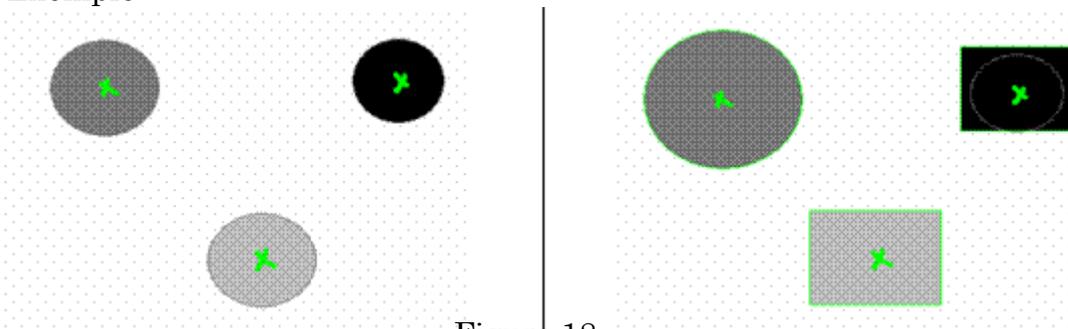


Figure 18.

Ces deux images ont les mêmes positions des centres de gravité de ses régions donc elles ont le même MST mais elles sont différentes. Pour cela on a décidé d'extraire d'autres signatures.

4.2. Deuxième Méthode : calcule des angles et des rapports entre les segments figurant dans une image. [3]

L'idée de cette méthode est d'exploiter les attributs géométriques et les informations structurels pour construire un Histogramme. En calculant les angles et les rapports entre les segments figurant dans une image, on capte la topologie de l'objet cible car pour un ensemble de segments donnés, ces angles et ces rapports sont uniques. Cette méthode prend les mêmes entrées que la méthode précédente.

4.2.1. Algorithme :

Entrées : image segmentée

Sorties : Histogramme.

Etape de l'algorithme :

- 1) Segmenter l'image
- 2) Extraire les axes d'inertie des régions.
- 3) Calculer des attributs géométriques :
 - a. Les angles entre les axes d'inertie deux à deux.

$$\alpha_{ab,cd} = \arccos \left[\frac{\mathbf{x}_{ab} \cdot \mathbf{x}_{cd}}{|\mathbf{x}_{ab}| |\mathbf{x}_{cd}|} \right];$$

- b. Le rapport selon la formule suivante :

$$\vartheta_{ab,cd} = \frac{1}{\frac{1}{2} + \frac{D_{ib}}{D_{ab}}}.$$

Figure 20.

2) Alors pour chaque couche d'information on calcule l'histogramme.

2.1. Extraire les axes d'inertie



Figure 21.

2.2. Calculer les angles et les rapports. Puis construire l'histogramme :

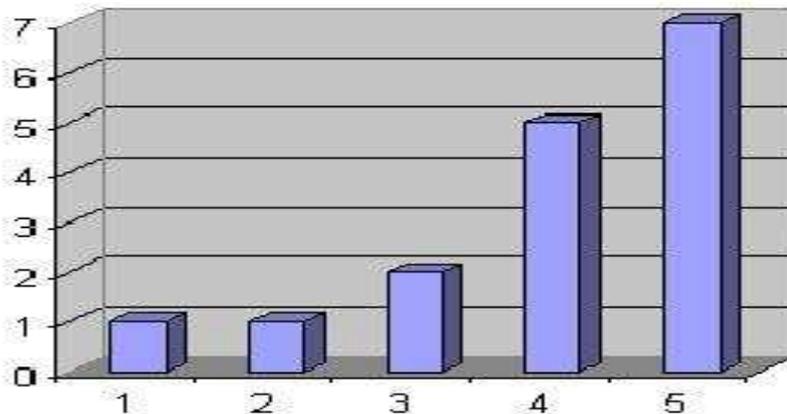


Figure 22

Alors pour chaque image on aura une signature :



Figure 23.

Dans le terme de mesure de similarité entre deux images, nous utilisons la distance de Bhattacharyya pour comparer deux histogrammes, dont sa formule est :

$$B(Q, M) = - \ln \sum_{I=1}^{TR} \sum_{J=1}^{TR} \sqrt{h_Q(I, J) \times h_D(I, J)}.$$

Avec :

n_A = le nombre des angles.

n_R = le nombre des rapports.

h_Q = l'histogramme de la première image.

h_D = l'histogramme de la deuxième image.

4.3. Conclusion

Dans ce chapitre, deux descripteurs de forme s'appuyant respectivement sur les arbres minimaux et les attributs géométriques sont proposées. Les algorithmes et les tests sont bien détaillés. En reliant tous les points caractérisant un objet, sous la contrainte de définir un chemin minimal et en calculant les angles et les rapports entre les segments figurant dans une image, une mesure de la topologie de l'objet est définie. Par construction, le MST réalise une mesure de cette topologie car pour un objet donné la longueur du MST est unique. Néanmoins, cette application n'est pas bijective, à une longueur donnée peut correspondre deux objets de formes différentes. Et un autre problème figure dans le calcul des angles c'est comment choisir l'orientation. Les premiers résultats testés sur des images de documents anciens issus de bibliothèques du centre d'études supérieures de la renaissance de Tours ont montré l'efficacité de ces descripteurs dans la recherche des images par le contenu.

Chapitre 5: Conclusion et perspective

Tout au long de ce rapport nous avons présenté les divers volets du travail réalisé pendant les six mois du Master 2 Informatique, Mathématiques et leurs Applications (IMA). Le problème abordé été la recherche d'une représentation haut-niveau pour l'indexation d'images par le contenu. Ce sujet est très intéressant. Souhaitant travailler dans le domaine de l'analyse d'image, ce stage a été très enrichissant pour moi. Pour ma part, pendant ces six mois, j'ai vécu des moments enrichissants. Dans un premier temps au niveau humain, les relations avec les personnels du laboratoire ont été très cordiales, j'ai trouvé en permanence du soutien et de la responsabilité pour répondre à mes questions. Je comprends mieux maintenant le but et le fonctionnement de la recherche. Dans un second temps, au niveau plus personnel, lors des échanges de point de vue, ce sont les propositions suggérées au cours des réunions hebdomadaires qui m'ont permis de découvrir et d'aboutir, après un approfondissement bibliographique, à trouver une solution aux différents problèmes qui apparaissaient au fur et à mesure de l'avancement du projet. Je termine ce stage avec la certitude d'avoir acquis des connaissances importantes. De plus le fait que ce stage soit parti intégrante du projet MADONNE, qui réunit plusieurs laboratoires apporte aussi une valorisation supplémentaire à mon travail.

Ce rapport commence par une présentation du sujet du stage et du laboratoire L3i. Puis dans le 3^{ème} chapitre rentre dans le problème de notre étude et propose un état de l'art des méthodes d'indexation d'image par le contenu. Les premières approches développées en indexation d'images calculaient ainsi des statistiques globales aux images pour les caractériser (par leur couleur ou leur texture par exemple). Depuis une dizaine d'années, des modes de représentation plus évolués sont employés. Ces méthodes dites « spatiales » prennent en compte l'arrangement des différents objets au sein des images. L'utilisation de cette information supplémentaire permet alors une caractérisation beaucoup plus fine des images.

Ayant montré la nécessité d'une représentation des images pour l'indexation, le 4^{ème} chapitre présente les deux nouvelles méthodes développées pour extraire des descripteurs utilisés dans la recherche des images par le contenu. La première méthode utilise la longueur du MST comme signature d'une image. La deuxième méthode calcule les angles et les rapports entre les segments figurant dans une image. Dans les deux cas on capte la topologie de l'objet cible car pour un ensemble de points donnés la longueur du MST est unique. De même les angles entre un ensemble de segments donnés sont uniques.

Ce travail prospectif fournit des résultats très prometteurs. Le premier sujet qui est sûrement le plus important pour améliorer la robustesse du MST c.à.d que la fonction de calcul de la longueur du MST soit bijective. Le deuxième sujet est de choisir la meilleure orientation adoptée pour le calcul des angles entre les segments figurant dans une image.

D'autres méthodes [4], [5], [6] peuvent être implémentées. Dans [6], la méthode est basée sur un algorithme de mise en correspondance de graphes, adapté aux graphes d'adjacence de

régions. Le nombre de régions représentant l'objet peut différer d'une image à l'autre. L'article propose d'une part une solution pour réduire la combinatoire de l'appariement de graphes et d'autre part une mesure de similarité entre graphes permettant d'ordonner les images résultats. L'adjacence et la relation spatiale entre les régions sont employées pour accélérer la recherche.

Bibliographie

- [1] : Thèse Julien BOMBRE, "Systèmes de représentation multi-echelles pour l'indexation et la restauration d'archives médiévales couleur ". Faculté des sciences Fondamentales et Appliquées, Université de Poitiers.
- [2] : Jean-Marc Ogier, Patrick Franco, "Proposition d'un modèle de description d'un objet dans une image". Laboratoire Informatique, Image, Interaction (L3i). Pôle Sciences et Technologies, Université de La Rochelle.
- [3] : B. Huet and E.R Hancock, "Line Pattern Retrieval Using Relational Histograms", IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Computer Society Press, vol. 21, no. 12, pp. 1363-1370, 1999.
- [4] : Nozha Boujemaa, Fauqueur, "Recherche d'images par régions d'intérêt : segmentation grossière rapide et description couleur fine". RSTI-TSI- 22/2003. Indexation par le contenu visuel, pages 1107 à 1138.
- [5] : Sylvie Philipp-Foliguet, Mohamed Lekkat, "Image retrieval from a partial request using region disposition ". Equipes Traitement des images et du signal, CNRS 8056, ENSEA/UCP, Cergy, France.
- [6] : Sylvie Philipp-Foliguet, Mohamed Lekkat, "Image retrieval from a partial request using region disposition ". Equipes Traitement des images et du signal, CNRS 8056, ENSEA/UCP, Cergy, France.