

# RECONNAISSANCE D'OBJETS EN IMAGERIE AERIENNE

*SADGAL Mohammed\*, EL FAZZIKI Aziz\* & AIT OUAHMAN Abdellah\*\**

*Faculté des sciences Semlalia Marrakech*

*Département Informatique\**

*Département de Physique\*\**

*Bd Moulay Abdellah*

*B.P 2390.*

*{sadgal,elfazziki, aitouahman}@ucam.ac.ma*

## 1. Introduction

L'analyse et l'interprétation d'images aériennes ou satellitaires constituent un champ important de recherche et d'études scientifiques. Plusieurs applications en tirent profit, par exemple la construction et la mise à jour des cartes topographiques, la mise à jour des plans cadastraux, la découverte des maladies des forêts, la localisation des zones de pollution, etc. Nous travaillons depuis 1999 dans ce domaine et nous portons nos efforts sur l'extraction de l'information, et spécialement sur la classification et la reconnaissance d'objets [Sad00][Sad01a][Sad01b][El K00][El K02a][El K02b]. La complexité de traitement exigé par ces applications a motivé plusieurs chercheurs en vision en s'impliquant dans la résolution des problèmes particuliers [El K99] [Gar99] [Wei95] [Wal98]. On trouve alors parmi ces problèmes: l'extraction des routes ou des réseaux routiers [Air95] qui est considérée comme une tâche importante pour la mise à jour des cartes dans une base de données. Cela nécessite non seulement la reconnaissance d'objets (routes) mais aussi une grande précision sur leur géométrie. Les travaux qui prennent en charge les routes montrent que les contrastes ne sont généralement pas assez forts pour employer des algorithmes de détection de contours. Les images nécessitent donc l'utilisation d'une segmentation en régions. On ne cherche pas des discontinuités mais au contraire des zones homogènes au sens d'un ou plusieurs critères donnés. Cette problématique caractérise les régions par leur luminance et par la texture. Par contre la détection des réseaux routiers se base sur les techniques de poursuite de contours et l'utilisation des connaissances telles que: forme, connexité, courbure. La localisation des forêts, des zones de végétations, des zones d'habitation, etc. prend une partie importante de traitement. L'extraction de ce type d'information fait jouer l'aspect texture comme

facteur indéniable. Le plus souvent on se base sur des techniques statistiques qui ont donné naissance à plusieurs approches dont l'essentiel est l'approche markovienne [Won92]. Mais il existe toute une panoplie d'approches traitant des zones homogènes [Che95].

Les formes planes n'ont vraiment de l'importance que dans la phase de reconnaissance où on cherche une zone particulière (par exemple une parcelle spécifique, une route, un jardin, etc.), ici la segmentation par détection des bords peut accomplir ce type de tâches. L'information tridimensionnelle peut concerner les montagnes, les buildings, etc. et traitée dans différentes méthodes décrites dans [Bas96][Bre92].

Dans ce type d'images, on aborde généralement deux tâches : Délimiter des zones dans l'image (Segmentation) et déterminer ce que peut être ces zones (Reconnaissance).

Bien que la littérature en traitement d'images soit abondante par les méthodes de **segmentation**, il n'existe pas d'approches générales. Il est plus judicieux de d'établir un environnement permettant la coopération de plusieurs méthodes.

**La reconnaissance d'objets** est un problème épineux se plaçant au niveau supérieur dans la hiérarchie des tâches de vision et constitue la partie computationnelle la plus difficile. Pour surmonter cette difficulté, un système de vision doit être capable de combiner ses capacités de représentation interne dans le but de réussir des décisions.

Les schémas computationnels actuels de reconnaissance se limitent généralement à l'utilisation d'un petit nombre de modèles spécifiques à un domaine d'intérêt donné. Ils n'abordent qu'un nombre restreint de types d'objets (souvent un seul type d'objets est pris en compte, par exemple: la reconnaissance de visages humains). Cette approche est utile pour réaliser des actions requérant une certaine précision. Elle devient cependant réductrice lorsque des scènes complexes nécessitent l'utilisation d'un grand nombre de modèles qui est le cas des images aériennes. Beaucoup de difficultés apparaissent, notamment celles liées à la variabilité d'apparence liée : à la lumière, à l'orientation, etc.

Ce papier traite ces deux aspects problématiques (segmentation et reconnaissance) dans le but d'améliorer la qualité du traitement et de minimiser le temps de réponse global d'un système de vision.

Nous avons mis en avant deux objectifs :

Le premier objectif de notre travail, est la mise en place d'une représentation interne d'objets facilitant la tâche des algorithmes de traitement d'images et essentiellement ceux d'apprentissage et de reconnaissance. Le but ici ne se limite pas uniquement à la détermination des modèles d'objets mais il s'étend sur une organisation efficace rendant possible la cohérence de plusieurs modèles à l'image du monde réel. Naturellement, le monde se constitue en catégories ce qui porte notre réflexion sur les manières selon lesquelles les objets se trouvent ainsi organisé.

Notre troisième objectif est l'introduction de la coopération comme mécanismes de recherche de qualité de traitement. Faire coopérer plusieurs méthodes pour une tâche ne peut qu'améliorer sa qualité. Par exemple pour une tâche de segmentation, une méthode basée sur la détection de contour fournit la forme d'une zone, alors qu'une méthode basée sur la texture fournit le contenu de cette même zone. La coopération permet d'avoir les deux informations en même temps. Bien entendu il faut vaincre la complexité computationnelle par l'accomplissement du premier et deuxième objectif. Notre but est de mettre en œuvre une architecture capable de fournir un environnement favorisant l'exécution de plusieurs tâches en parallèle d'une part et d'améliorer le rendement des algorithmes d'apprentissage et de reconnaissance d'autre part. Cette amélioration passe par la mise sur pied d'une stratégie de recherche optimale se basant sur l'organisation évoquée dans le premier objectif.

Pour atteindre ces objectifs, nous avons introduit la catégorisation comme un choix fondé sur le besoin d'organiser le monde d'objets (ici au sens large: objets physiques, événements, idées,...) constituant la base des systèmes de vision et d'établir un environnement propice pour les procédures de calcul dans le but de réduire la complexité computationnelle (temps, espace mémoire,...). Toutes les opérations pertinentes à la vision numérique requièrent des ressources de calcul intensives, une attention particulière doit être portée sur l'analyse de la complexité des algorithmes de vision et sur la possibilité de les paralléliser ou des les adapter à des architectures spécialisées. Notre architecture se ressource de l'idée qui favorise « la parallélisation » et la coopération des tâches.

L'apport de ce travail réside essentiellement dans la création d'un environnement permettant la coopération entre plusieurs méthodes au niveau de la segmentation,

l'apprentissage et la reconnaissance d'une part et l'indexation de la base d'objets par la catégorisation pour l'optimisation des tâches de la reconnaissance d'autre part.

Nous aborderons ces problèmes dans les paragraphes 2 et 3. Nous proposons ensuite une démarche permettant d'épauler la reconnaissance. Il s'agit d'organiser les systèmes de vision pour que le processus de la reconnaissance démarre à partir de la segmentation dans le but d'optimiser le nombre d'objets à la fin de ce processus. Le paragraphe 4 présente un exemple illustrant différentes opérations de la segmentation multicritères et catégorisation.

## **2. Traitement d'images : différentes approches**

La plupart des théories de la reconnaissance d'objets ne traitent que l'aspect géométrique (forme 3D des objets). On peut compter aujourd'hui deux grands groupes de théories qui divergent sur le format de la représentation selon que celui-ci est indépendant ou dépendant des vues de l'objet à représenter.

Le premier groupe de ces théories considère que la représentation d'un objet est conçue comme un ensemble de caractéristiques (invariants) de l'objet qui sont indépendantes des vues de ce même objet [Wal97]. Il s'agit d'une description structurelle de l'objet. L'une des approches les plus intéressantes et celle de Biederman : «Reconnaissance Par les Composantes » (RPC) [Bie87] qui consiste à représenter l'objet en le décomposant en des structures (primitives : geons ) selon un schéma proposé par de Marr et Nishihara [Mar78].

Le second groupe de ces théories considère que la représentation d'un objet est liée à des vues spécifiques à l'objet et que n'importe quelle autre vue pourra être déduite à l'aide de ces vues[Ede99][Tar98a]. Les modèles de ce type de représentation considère une vue comme une collections de caractéristiques (informations 2D, information 3D,...). La reconnaissance est exprimée comme une fonction des images déjà vues. Plusieurs modèles computationnels utilisés notamment par Edelman et Poggio basé sur des expériences [Pog90][Tar98b].

Le modèle revendiqué par le premier groupe peut apparaître attractif grâce à une représentation compacte et robuste pour les objets. Or l'expérience a montré qu'il est très difficile de détecter les invariants dans les images d'une part et que cette structure

d'invariants conduit forcément à la catégorisation et non pas à l'identification d'autre part.

Pour remédier à ces inconvénients, des nouvelles approches ont fait surface selon la théorie du deuxième groupe. Ces approches modélisent les objets par leurs images elles-mêmes délaissant ainsi les modèles de type géométrique ou fondés sur des structures d'invariants. Pour cela, la sphère des vues possibles est échantillonnée et chaque élément de cet échantillonnage est une des apparences de l'objet. Ainsi, un objet est représenté par une collection d'images et la reconnaissance est basée sur l'appariement d'une nouvelle image de l'objet avec les images de la collection. Plusieurs travaux [Ede99][Rie00] ont évolué dans ce sens.

## **2.1. Segmentation**

En analyse d'images, on fait souvent la distinction entre traitements *bas niveau* et *haut niveau*. Les premiers travaillent sur les valeurs attachés aux pixels de l'image sans faire le lien entre la réalité qu'elles représentent. Tandis que les seconds opèrent sur des entités symboliques constituant une interprétation de la réalité extraite de l'image. La segmentation est un traitement bas niveau qui consiste à diviser l'ensemble de pixels en régions connexes, homogènes et différentes de ses voisins. Ici, on ne cherche pas à déterminer ce que les régions représentent. La qualité de l'interprétation d'une image dépend fortement de celle de la segmentation. Malgré la grande diversité de méthodes, les résultats de segmentation restent moyens et varient beaucoup en fonction de la technique choisie. Une méthode de segmentation générale et automatique est difficile à concevoir étant donné les différents types de régions pouvant être présentes dans une image.

Généralement, la segmentation d'une image est effectuée par l'utilisation de l'une des deux grandes approches basée sur l'extraction de contours (frontières) ou la croissance des régions [Tup98][Lop99][Wro87].

**Plusieurs techniques sont utilisées :**

- ***l'approche contour*** : L'approche contour consiste à identifier les transitions entre les régions. L'approche contour n'aboutit pas directement à une segmentation, car les contours détectés ne sont pas toujours connexes. Il existe cependant des techniques permettant d'obtenir des contours fermés. Dans ce cas, on observe une parfaite dualité entre les contours et les régions.

Dans cette approche la détection de contours consiste à balayer l'image avec une fenêtre définissant la zone d'intérêt. A chaque position, un opérateur est appliqué sur les pixels de la fenêtre afin d'estimer s'il y a une transition significative au niveau de l'attribut choisi. Littérature est très riche par ces techniques : [Can86][She96] [Der88].

- ***L'approche région*** : Les méthodes de l'approche région cherchent à regrouper directement des pixels ayant une propriété commune; l'ensemble des regroupements de pixels définit à la fin une segmentation de l'image. Les plus importantes sont celles qui procèdent par croissance de régions et par division/fusion de régions :

La segmentation par croissance de régions [Lop99][Cha94] vise à regrouper les pixels adjacents de l'image dont les attributs varient de façon négligeable. Il faut choisir un prédicat ou critère d'uniformité P et un jeu de pixels initiaux, ou germes, à partir desquels la croissance de régions s'effectuera.

Les méthodes du type division et fusion comportent, comme leur nom l'indique, deux étapes. Dans la première, l'image est divisée récursivement jusqu'à ce que toutes les régions vérifient le prédicat P. Dans la deuxième étape, des régions adjacentes sont regroupées, tant que les régions résultantes vérifient P.

- ***Coopération*** : Concernant la segmentation, il est acquis depuis une dizaine d'années déjà qu'il faut faire coopérer les approches régions et les approches contours [Wro87] un bon critère de segmentation. Indépendamment des problèmes inverses à régulariser, de nombreux travaux ont été développés ces dernières années sur des modèles de contours. Ces modèles de contours sont développés pour la segmentation d'objet dans une image et n'étaient pas, jusqu'à récemment encore, envisagés comme des modèles de contours pour la segmentation par approche contour/région. Les travaux actuels tentent de combiner et d'unifier ces différents modèles.

## **2.2. Classification**

Contrairement à la segmentation d'images, ici on cherche à identifier ce que représente chacune des régions segmentées. En imagerie aérienne par exemple, on cherche à déterminer les zones urbaines, les forêts, zones d'eau, les montagnes, ....

Il s'agit de regrouper les différents éléments (pixels) en thèmes correspondants à la vérité du terrain. Le résultat est bien entendu une image segmentée. Dans ce domaine on procède, par l'attribution des pixels de l'image à des classes connues a priori (c'est la classification supervisée) ou à des classes inconnues (classification non supervisée).

Plusieurs approches sont utilisées à ce sujet (nous décrivons les principes de certaines d'entre elles):

- **Minimisation de distance** : Le principe des méthodes reposant sur la minimisation de distance consiste à rechercher la classe la plus proche pour chaque pixel, ou groupe de pixels si l'on travaille dans une fenêtre d'analyse centrée sur le pixel courant. La notion de proximité est liée à la distance considérée.

- **Les approches stochastiques** : (*Champs de Markov, inférence bayésienne, règle de Bayes*)

Dans ce type d'approches, les modèles sont souvent discrets et de nombreux travaux reposent sur la théorie des champs de Markov et l'estimation bayésienne [Ber96][Des99].

Les modèles markoviens en traitement d'images [Won92][Che95] ont trouvé leur application dans divers problèmes de vision par ordinateur comme : segmentation, classification, restauration d'image, stéréovision, imagerie radar, modélisation de textures, .... Les premières approches proposées ont fait appel à des techniques de relaxation stochastique fondées sur le recuit simulé.

- **Les Réseaux de neurones** : Les réseaux de neurones [Wid90] s'avèrent capables de traiter des problèmes complexes de reconnaissance de forme, ou de simulation de processus non linéaires et/ou dynamiques. Ils sont souvent utilisés pour leur capacité de classifieur et ils sont classiquement employés dans des problèmes d'approximation pour simuler des fonctions de transfert non-linéaires et multi-dimensionnelles et pour résoudre des problèmes d'inversion [Cha00]

#### **2.4.Reconnaissance d'objets (l'aspect général)**

La reconnaissance d'objets est un domaine très actif en vision par ordinateur. L'approche générale consiste à considérer que l'on dispose d'une banque de données où sont stockés des modèles d'objets et une vue du monde réel. On doit répondre à une question essentielle : L'objet extrait (par les moyens précédents) existe-t-il dans la banque de données. Le problème dépend de la modélisation d'objets adoptée. Dans la majorité des cas cette modélisation est purement géométrique : un ensemble de caractéristiques appelées primitives. L'opération de reconnaissance consiste en l'appariement des ces caractéristiques et celles décelées dans l'image.

On utilise des techniques d'alignement qui consistent à transformer le repère des modèles en celui des données (image). Des travaux significatifs dans ce sens ont été présentés par Ayache et Faugeras [Aya86] (système HYPER), Lowie [Low87] (système SCERPO),...

### **3. Une architecture pour les systèmes de traitement d'images aériennes**

La catégorisation constitue la pièce maîtresse de notre architecture, car quelque soit la manière dont on mémorise les objets, les événements, les idées, ..., leur différenciation et leur regroupement semble un processus actif permettant d'apprendre et de reconnaître. Ce processus aboutit naturellement à une organisation du monde en catégories. Il est donc un moyen capable de fournir une représentation interne d'une part et de permettre une opération d'indexage sur les objets mémorisés d'autre part. Cet indexage est un élément important dans l'optimisation de recherche d'information.

On se base sur deux principes pour la mise en pratique du processus de catégorisation : principe d'économie, dans ce cas

il doit équilibrer entre deux tendances opposées. La première tendance consiste à utiliser les catégories de manière à maximiser la quantité d'informations qu'elles nous donnent. Ceci conduit à former le plus grand nombre de catégories possibles, puisqu'une catégorie plus spécifique contiendra plus d'informations sur ses membres. Plus on a de catégories, plus on peut faire de distinctions entre objets et événements. La tendance opposée qui nous pousse à avoir des catégories est la tendance à réduire la quantité d'information que nous avons à traiter pour la rendre cognitivement utilisable. Un organisme a avantage à ne pas faire de différence entre des objets quand ces différences sont sans pertinence pour ce qu'il veut faire. Le degré de finesse de nos catégories va donc être le résultat d'un compromis entre ces deux tendances.

Le deuxième principe qui gouverne l'organisation de nos catégories est le principe d'exploitation de la structure du monde perçu [Rosh78].

Nous présentons ce paragraphe suivant quelques aspects de la catégorisation qui ont guidé le choix de notre modèle.

#### **3.1. Catégorisation : Définitions, Modèles et Critiques**

##### **3.1.1. Catégorie**

Classiquement, une catégorie est vue comme une collection d'instances qui peuvent être traitées de la même façon indifféremment. Cette collection peut être constituée de : objets, événements, idées, etc.

De cette manière classique, les membres d'une catégorie sont déterminés par l'appariement d'une définition : les membres sont définis à l'aide d'un certain nombre de caractéristiques qui peuvent être totalement ou partiellement nécessaires. *Exemple : Un triangle est une surface avec 3 cotés et 3 sommets.*

Cette façon de catégoriser les membres pose des problèmes :

- Il est très difficile de spécifier un ensemble complet des caractéristiques (nécessaires et suffisantes), par exemple : définition de la catégorie **Jeu**.
  - Les exemples d'une catégorie naturelle ne sont pas jugés de la même façon : certains sont plus typiques que d'autres. (même des définitions formelles : 3 est plus typique que 59 pour la catégorie des nombres impairs)
  - Ambiguïté : des membres peuvent se trouver dans plusieurs autres catégories
- **Modèle des prototypes** : Rosch [Ros76][Ros78] suppose que, à partir d'expériences avec des exemplaires, se forme et se développe une « tendance centrale » d'une catégorie et que le jugement d'appartenance à une catégorie se base sur cette « tendance centrale » ou « prototype ». Le prototype est ainsi défini comme une représentation abstraite possédant les caractéristiques moyennes des exemplaires pouvant appartenir à une catégorie donnée.

L'idée de représenter une catégorie par un prototype est due à plusieurs raisons :

- Avec ce modèle, il n'est pas nécessaire de donner la liste complète des caractéristiques indispensables et suffisantes. Les catégories sont définies par un ensemble de caractéristiques les plus typiques.
- Les membres partagent une famille de ressemblances.
- L'évidence de la typicalité : les instances prototypiques sont catégorisées rapidement, les membres prototypiques sont listés et appris en premier et ils sont utilisés comme des points de référence.

Il y a deux façons de définir un prototype :

- a) Un prototype est le membre le plus typique du groupe.
- b) Le prototype peut être considéré comme une moyenne de tous les membres du groupe ( Il ne correspond à aucun membre du groupe ).

- **Modèle des exemplaires** : Comme l'ont noté Murphy & Medin [Mur85], les principaux défauts de la théorie des prototypes sont sans doute liés à l'insuffisance de l'idée que les catégories sont représentées par des listes de traits et que la catégorisation est fonction du degré de similitude avec le prototype catégoriel. On doit noter tout d'abord un défaut de contraintes: la théorie des prototypes n'explique pas ce qui fait qu'un trait est ou non pertinent et ne nous dit pas comment définir la similitude. Les relations de similitude entre un groupe d'entités dépendent dans une large mesure de l'importance, du poids, que l'on accorde aux attributs individuels. Pour résoudre le problème de l'influence du contexte, et celui de la variabilité, on stocke toute instance (appelée exemplaire [Eri98]) de la catégorie. Deux autres avantages : il est plus facile de générer un exemplaire qu'un prototype, et la variation des caractéristiques des exemplaires est constatée. La détermination d'un membre de la catégorie se fait par similarité selon la règle du plus proche voisin (on tient compte de tous les exemplaires). La difficulté ici est la complexité de calcul.

- **Modèle hybride** : L'hybridation ici consiste en l'utilisation de deux modes de traitement des données. Le traitement dirigé par données est déterminé par l'information-stimulus venant de l'extérieur, alors que le traitement dirigé par concepts est guidé par les informations déjà stockées en mémoire, ce qui implique une expérience préalable et des concepts déjà acquis. Dans le cas du traitement dirigé par données, on parle de processus "bottom/up" (l'interprétation du stimulus se faisant de "bas en haut", c'est-à-dire par interprétations successives jusqu'au résultat final). On parle de processus "top/down" dans le cas du traitement dirigé par concepts qui implique la génération d'hypothèses. Le modèle de Neisser [Nei67] incorpore les deux traitements en soulignant que leur importance dépend du type de stimulus et de l'information contextuelle.

## **3.2. Description de l'approche**

### **3.2.1. Le modèle :**

Compte tenu de la particularité des objets de notre domaine d'application (images aériennes), il faut enrichir la représentation interne par d'autres aspects comme la couleur et la texture et doter le système d'un mécanisme d'inférences pour palier les lacunes du modèle des prototypes.

Dans ce travail nous considérons que la catégorisation est un moyen efficace pour organiser un système de vision. La représentation interne doit tenir compte des niveaux d'abstraction pour fournir une reconnaissance graduelle.

Nous retenons des travaux de Rosch sur le niveau de base deux enseignements intéressants :

- 1- A l'intérieur d'une catégorie, on doit avoir le maximum de ressemblances entre les membres (prédiction)
- 2- Entre les catégories, on doit avoir le maximum de différences. On veut avoir des catégories distinctives.

A cela nous supposons que le niveau supérieur est un niveau organisateur et générateur comprenant des connaissances sur la validité (conformité) d'une catégorie.

Notre modèle sera donc à prototypes avec l'introduction des connaissances au niveau supérieur.

Nous organisons les caractéristiques sous une forme hiérarchique en trois niveaux :

- Le premier niveau concernera des informations telles que *la couleur, la texture, les formes planes, etc.*
- Le deuxième niveau se concentrera sur *les parties d'un objet*. La reconnaissance d'une partie peut sélectionner tous les prototypes contenant cette partie.
- Le troisième niveau s'intéressera à des *contraintes* intra-objets et inter-objets [Sad99]. Une contrainte peut suggérer un ensemble de prototypes capable de la valider.

Le fait d'accompagner les prototypes par plusieurs types d'information permet donc une indexation implicite des images (vues) concernant les objets stockés dans la base de données.

Pour cela, il faut doter l'objet de la représentation la plus complète possible selon les aspects suivants:

- 1) la couleur
- 2) la texture

- 3) les formes planes (surfaces).
- 4) Le volume (3 dimensions)
- 5) La décomposition de l'objet en ses parties.

Pour illustrer notre idée, on va s'intéresser au premier niveau de traitement et prendre en compte la couleur, la texture et les formes planes (2 dimensions). Ces caractéristiques sont déduites dans une structure prédéfinie de paramètres ou à partir d'une série de tests qui consiste à présenter au système d'un ensemble d'images du même objet (dans une phase d'apprentissage) ou introduites directement (manuellement) dans cette structure. Il s'agit de mémoriser les caractéristiques qui sont partagées par les images du test (c'est-à-dire les paramètres des couleurs pour la végétation et ceux de textures particulières pour les arbres, etc.). Ensuite ces caractéristiques sont comparées à celles des objets les plus proches pour générer la catégorie résultante (considérée ici comme catégorie de base).

La modélisation prend en compte les trois niveaux d'abstraction : sur\_catégorie, base\_catégorie et sous\_catégorie selon le schéma suivant:

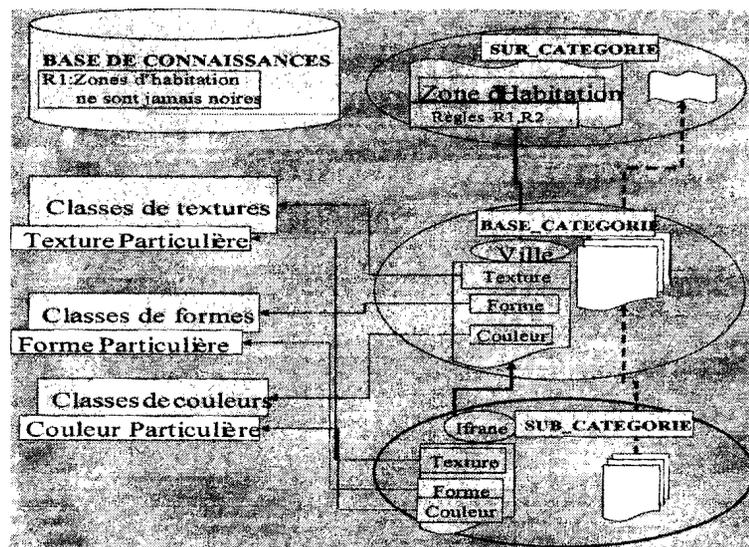


Figure 1 : ( Ifrane : une ville marocaine)

Ce schéma explicite l'organisation des modèles d'objets selon les trois niveaux de catégorisation. Nous avons introduit en plus l'aspect connaissances que peut renfermer les catégories : règles de conformité, procédures de déductions, etc. Nous insistons ici sur l'aspect visuel, les caractéristiques peuvent être introduites par des paramètres illustrés ici dans la partie gauche. Chaque caractéristique pointe vers un ensemble de paramètres selon sa position (SUB\_CATEGORIE, BASE\_CATEGORIE OU SUR\_CATEGORIE): Les caractéristique en SUB-CATEGORIRE pointent (pointeurs rouges) vers des éléments particuliers par exemple couleur verte. Par Contre les caractéristiques de BASE\_CATEGORIE pointent (pointeurs bleus) vers des classes d'éléments par exemple classe des triangles. Pour cette raison, les paramètres sont organisés, à leur tour, par catégories : couleurs, textures, formes, etc.

La mise en relief des ces trois niveau est naturelle. D'abord le niveau de base qui est le niveau d'accès par excellence, car c'est effectivement l'aspect visuel qu'on garde à ce niveau. Les détails sont exprimés au niveau SUB\_CATEGORIE, cela est nécessaire lorsqu'il s'agit d'une identification (par exemple le visage d'une personne particulière). Le contenu du niveau supérieur (SUR\_CATEGORIE) exprime les contextes et il n'est pas figé et évolue selon l'expérience.

Les caractéristiques mentionnées (couleurs, texture, forme, etc.) dans ce schéma ne sont qu'à titre indicatif. Nous verrons plus loin (partie architecture) qu'il y a plusieurs façons de modéliser les objets.

Nous traitons des images aériennes où il faut reconnaître des zones d'habitation, des forêts, des régions montagneuses, des routes, des rivières, etc. En imagerie urbaine par exemple, et en raison de la résolution, de nombreux détails ne sont pas visibles de façon nette, les voitures, les arbres isolés et tous les autres petits objets.

Les principales difficultés [Bai97] présentées par ces images proviennent de :

1. **Diversité des objets** : Le milieu urbain regroupe des objets très différents : bâtiments pavillonnaires ou industriels, murets, végétation plus ou moins étendue, etc.
2. **Complexité** : Les objets impliqués peuvent avoir des formes arbitrairement complexes.
3. **Densité** : la densité des objets du sursol est l'origine de nombreuses interactions : adjacences, parties cachées, ombres, etc.

Il est donc très difficile de se baser sur un seul critère de segmentation. En effet, les régions homogènes sont très souvent soumises à la sous ou sur-segmentation. De même, l'extraction des primitives de type contour est souvent faussée par les adjacences : chaînage non cohérent avec la réalité de la scène, parties linéaires morcelées ou déviées, le paramétrage est alors délicat [Cor97]. D'où l'intérêt de mettre en compétition une segmentation à plusieurs critères.

Exemple de photos aériennes – basse altitude- (figure 2):



**Figure 2 :** Site Vallée de Dades (Maroc)

### ***3-2-2. L'architecture générale :***

Notre architecture est basée sur l'idée qui consiste à considérer la reconnaissance comme un processus graduel aboutissant à une identification ou à une catégorisation. Elle possède trois grands modules (figure 3a): Une bibliothèque contenant les outils de traitement d'images, un module concernant la segmentation multicritère (figure 3b) et un module pour les tâches de catégorisation (figure 3c).

Le processus de reconnaissance proposé dans cette architecture suit une démarche simple:

- 1- Mise en œuvre de la segmentation multicritère pour l'extraction la plus complète possible des indices (caractéristiques de texture, forme et couleur)
- 2- Sélection des catégories au fur et à mesure de la découverte des indices.
- 3- Elaboration de la segmentation par combinaison de critères et réduction du nombre de catégories par apport de précisions.
- 4- Décision final : Identification ou catégorisation

Le caractère essentiel de cette architecture est d'organiser la reconnaissance sous une forme permettant le déclenchement du processus depuis le niveau bas du traitement

d'images jusqu'au niveau le plus haut. Nous tentons alors, de résoudre deux problèmes : l'extraction d'objets via la segmentation et la reconnaissance des objets extraits.

Une image aérienne, dans notre cas, présente plusieurs objets, il serait très difficile de les extraire directement par le module de la segmentation. En effet, un objet pourrait être le résultat de plusieurs zones fournies par la segmentation multicritère. Il faudra alors un modèle d'objets efficace tenant compte de différentes caractéristiques. Pour cela, plusieurs approches mettant en avant des modélisations automatiques par l'apprentissage ont été avancées. Nous en citons quelques unes:

### 1- Modélisation par la couleur

Swain et Ballard [Swa91] ont proposé une technique selon laquelle chaque objet est représenté par un histogramme de couleurs (c'est à dire par une approximation de la distribution de ses couleurs). Leur technique est remarquablement fiable par rapport aux changements d'orientation d'objet, aux variations d'échelle, aux occultations partielles et aux changements de point de vue. Même des changements de la forme d'un objet ne dégradent pas forcément la performance de leur méthode. L'inconvénient principal de leur méthode est sa sensibilité à la couleur et à l'intensité de l'éclairage.

Plusieurs auteurs ont combiné l'approche des histogrammes de couleurs avec différents types d'informations géométriques pour la reconnaissance. SEEMORE est un des exemples utilisant les histogrammes de couleurs inspiré par des réseaux de neurones [Mel97].

### 2- Approches d'images propres

De nombreux chercheurs [Ohb96][Hal98][Art02] utilisent la transformation de Karhunen-Loeve [Fuk90] (ou analyse en composantes principales) pour le calcul d'images propres dans le contexte de la reconnaissance d'objets.

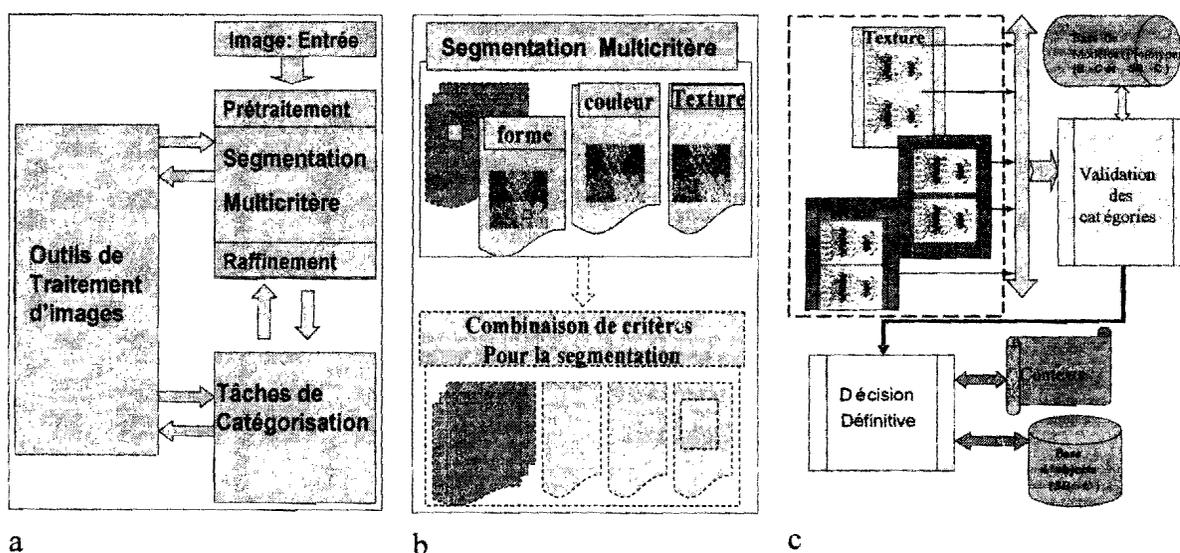
La transformation de Karhunen-Loeve est une technique classique de reconnaissance des formes. Néanmoins, elle est une des meilleures approches récentes de reconnaissance d'objets.

### 3- Densité probabilistique

Hornegger et Niemann [Hor95][Hor99] proposent une représentation d'objets par une densité probabiliste de caractéristiques d'objets. L'apprentissage et la localisation d'objets sont formulés comme des problèmes d'estimation de paramètres. Dans leur

formulation, les caractéristiques sont interprétées comme des variables aléatoires. Un objet est alors représenté par une densité probabiliste d'un ensemble de caractéristiques. [Hor99] utilise les coordonnées de points comme caractéristiques. Chaque transformation d'un objet (par exemple rotation, translation, changement d'échelle, changement de point de vue) est représentée par un paramètre de la densité probabiliste. Les auteurs s'intéressent, en particulier, au problème de la reconnaissance d'objets 3D à partir d'images 2D, et au problème de l'estimation d'un modèle 3D à partir des images 2D : Pendant l'apprentissage non supervisé (sans supposition de correspondance), la densité probabiliste d'un objet 3D est estimée à partir des images 2D en utilisant l'algorithme de maximisation de l'espérance.

Plusieurs autres méthodes ont été utilisées comme les Réseaux de neurones (déjà citée dans la paragraphe 3.3.1) et les programmes génétiques.



**Figure 3 :** a) L'architecture générale  
 b) Segmentation Multicritère c) Tâches de la Catégorisation

La modélisation automatique d'objets par des distributions de probabilités a apporté ses fruits dans beaucoup d'applications. Nous l'avons adopté ici pour illustration, mais la désignation d'autres tâches d'apprentissage et de reconnaissance comme le réseaux de neurones par exemples, ne remettra pas en cause le module de la catégorisation (figure 3c)

Comme nous l'avons précisé dans le paragraphe 3.2.1, nous traitons dans un premier temps les informations concernant la couleur, la texture et la forme. Nous combinons alors ces trois caractéristiques pour modéliser les objets par des 3 distributions de probabilités.

Le schéma ( figure 3 c) suggère un fonctionnement en deux modes : Apprentissage et Reconnaissance. Ici, l'apprentissage est pratiqué pour déterminer des modèles de textures, de couleurs et de formes. Le modèle d'objet (prototype) est formé par ces trois modèles.

La reconnaissance se situe au niveau des zones mais pas au niveau des pixels comme la classification.

Pour illustrer ces processus, on se place dans le cadre de la théorie Bayésienne. L'attribution d'un élément  $x$  à l'un des modèles  $\{\omega_1, \omega_2, \dots, \omega_k\}$  repose sur la formule suivante (formule de Bayes) :

$$\underline{\underline{p(w_i | x) = \frac{p(x | w_i).p(w_i)}{p(x)}}}$$

$p(w_i | x)$  est la probabilité *à posteriori* que l'élément  $x$  appartienne à la  $i^{eme}$  modèle. Pour un  $x$  donné, la reconnaissance consiste à l'attribuer au modèle qui maximise la probabilité *à posteriori* (on appelle cette valeur la *MAP estimate*).

$p(w_i)$  représente la probabilité qu'un élément pris au hasard appartienne à la  $i^{eme}$  modèle.

$p(x)$  est la loi modélisant l'histogramme global de l'image observée. On prend souvent une loi uniforme.

$p(x | w_i)$  est la probabilité *à priori* d'un élément  $x$  de la  $i^{eme}$  modèle. Les paramètres pouvant être déterminés automatiquement (classification non-supervisée), ou semi-automatiquement (classification supervisée).

**- Mode d'apprentissage:**

L'élément  $x$  est le pixel, le processus consiste à trouver des modèles de textures, de couleurs et formes.

Il s'agit ici de déterminer les différents paramètres des caractéristiques retenues (couleur, texture et forme) à l'aide d'un échantillon d'exemples d'images d'objets. On cherche alors les paramètres qui maximisent la loi de probabilité *a priori*  $p(x | w_i)$ .

**- Mode de reconnaissance:**

L'élément  $x$  est la zone et il s'agit de l'attribuer à un modèle d'objet connu qui rend maximal la probabilité *a posteriori*  $p(w_i | x)$ .

L'image brute est segmentée à l'aide de plusieurs critères : textures, couleurs, frontières,.... Selon chaque critère l'image ainsi segmentée présente des zones à l'entrée d'un processus de catégorisation qui sélectionne dans le cas échéant, un ou plusieurs prototypes (modèles). En effet, les groupements de caractéristiques provenant des régions segmentées et arrivant en premier dans le flot de données vont permettre d'indexer (présélectionner) un nombre réduit de prototypes compatibles avec ces données. Ces prototypes vont constituer des hypothèses initiales, qui devront être vérifiées par la présence dans l'image d'autres caractéristiques non encore détectées appartenant à ces mêmes objets. Sans ces mécanismes de réduction de données, le problème de la reconnaissance d'objets conduirait à une explosion combinatoire du nombre de possibilités à évaluer.

Pour affiner la catégorisation les zones peuvent être combinées (segmentation multicritères : par exemple combiner la texture et la couleur) ce qui donne de nouvelles zones plus précises (dans le sens de réduction du nombre de prototypes). La même stratégie est poursuivie dans la phase de l'identification, il suffit d'opérer par le même mécanisme de catégorisation sur les extensions (les objets particuliers dans la base) des prototypes sélectionnés dans les phases précédentes.

Ce processus est graduel, comme nous l'avons suggéré, commence par un ensemble de catégories possibles dès la phase de la segmentation, et tente de réduire cet ensemble au fur et à mesure à l'aide des détails découvertes dans l'image.

La partie « décision finale » consiste à procéder par l'appariement proprement dit avec l'utilisation du contexte, certaines hypothèses peuvent être éliminées.

#### 4. Exemple

Afin d'illustrer l'intérêt de la catégorisation dans un système de vision nous considérons l'exemple de l'image suivante à traiter (terrain nu, végétation et une forêt : figure 2 : vallée Dades )

On doit disposer d'une base de modèles pour les objets manipulés par exemple (figure 5)

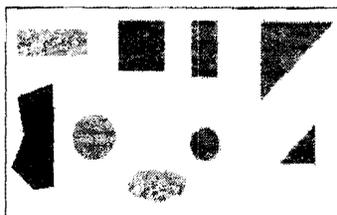


Figure 4-5

En subdivisant cette base selon la forme, la texture et la couleur, nous obtenons :

- trois prototypes de forme : rectangle, triangle, cercle et un trapezoide (PF1, PF2, PF3 et PF4)
- deux prototypes de texture: forêt et terrain (PT1 et PT2)
- quatre prototypes de couleur : vert, marron, bleu et rose (PC1, PC2, PC3 et PC4)

##### 4.1. Segmentation :

**Nous avons choisi 3 critères de segmentation pour cet exemples : frontières , couleur et texture :**

Le choix des algorithmes utilisés pour cet exemple est basé sur leur résultat intéressant dans la littérature.

Nous ne discutons pas ici ni leurs avantages et ni leurs inconvénients, car l'objectif est de montrer le niveau coopération entre différents critères de segmentation dans la phase d'apprentissage et de reconnaissance et non pas la comparaison de différentes méthodes pour chacun des critères.

- **Segmentation par la Détection des frontières :** La détection des frontières est basée sur les techniques de dérivation. Si  $g(x,y)$  représente le niveau de gris au pixel

(x,y) la quantité  $\rho(x, y) = \sqrt{\left(\frac{\partial g(x, y)}{\partial x}\right)^2 + \left(\frac{\partial g(x, y)}{\partial y}\right)^2}$  est le module du gradient en (x,y). Un point sur la frontière entre deux régions possède un saut important de niveau de gris. Le module du gradient en ce point doit être maximal selon la direction :

$$\theta(x, y) = \text{Arctg} \frac{\partial g(x, y) / \partial y}{\partial g(x, y) / \partial x}$$

En général, on procède à la segmentation par seuillage: les pixels retenus comme points frontières si leur module du gradient dépasse un certain seuil.

Sur le plan de calcul, on utilise des filtres de dérivation convolués à g(x,y) sous leur aspect discret. On cite ici les filtres simples les plus utilisés : **Sobel**, **perwitt**, **Kirch**, etc. L'opérateur **Sobel** selon les deux directions (horizontale et verticale sur une fenêtre 3x3) :

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Cet opérateur estime le gradient au point de coordonnées (x,y), en privilégiant les pixels proches du milieu du filtre. L'opérateur de **Sobel** est l'opérateur le plus communément utilisé dans la détermination de contours.

Pour améliorer la détection des contours nous avons adopté l'algorithme de Canny-Deriche qui utilise les cartes de gradient de l'image. Il permet de satisfaire les trois conditions lors de la conception des filtres de détection :

- 1) qualité de la détection des vrais points de contour,
- 2) localisation des points détectés,
- 3) unicité de la solution (en cas de détection multiple).

La prise en compte de ces trois critères pour l'opérateur idéal de détection a conduit Canny à proposer une première approximation par la dérivée de la fonction de Gauss (expression pour une dimension : 1D):

$$g_1(i) = \alpha . i . e^{-\frac{i^2}{2\sigma^2}}$$

Cette proposition a été améliorée par le filtre de Deriche par l'utilisation du filtre:

$$g_2(i) = \alpha . i . e^{-\alpha|i|}$$

Ce filtre a l'avantage de la séparabilité, car il opère indépendamment sur les deux axes de l'image.

L'implémentation du filtre de Canny utilise le lissage 1D avec le filtre de Gauss  $g(x)$ , suivi par la convolution avec le filtre 1D de la dérivée  $g_1(x)$  de la fonction de Gauss, sur chacun des deux axes, afin d'obtenir les cartes de la norme des gradients et de leurs orientations locales :

1. Filtrage de l'image par un filtre Gaussien dont la variance fixe le degré de lissage,
2. Estimation de l'amplitude et de l'orientation du gradient en appliquant un filtre de Sobel,
3. Sélection des maxima locaux,
4. Utilisation du seuillage par hystérésis pour détecter et relier les contours.

Les paramètres de l'algorithme de seuillage par hystérésis (consiste à appliquer deux seuils  $s_1$  et  $s_2=2*s_1$  et compléter le résultat de  $s_2$  par les données de  $s_1$ )  $s_1$  et  $s_2$ , sont choisis pour optimiser la relation entre la qualité de la détection locale et la qualité de la détection globale.

#### - **Segmentation par les Couleurs :**

Nous avons utilisé un algorithme efficace proposé et décrit par D. Comaniciu dans [Com99]. Il est basé sur l'analyse d'espace de caractéristiques. Dans ce paradigme les pixels sont groupés et exprimés dans un espace de couleurs. Chaque groupe représente une région homogène dans l'image. L'algorithme utilise une procédure appelée « Shift Mean » ou déplacement de la moyenne pour palier aux problèmes issus des techniques de « clustering » ( ) Soit  $x$  vecteur dans un espace de  $p$  caractéristiques et  $f(x)$  une fonction de densité de probabilités. Donnons nous un ensemble de point  $x_i$  dans le même espace, la propriété la plus simple de « shift Mean » peut être écrite comme suit :

$$\nabla f(x) = (\text{moyenne}(x_i) - x)$$

$$x_i \in S_r, x$$

où  $Sr, x$  est une hypersphère dans l'espace de  $p$  dimensions de rayon  $r$  et de centre  $x$ .  
Une application récursive de cette propriété permet de trouver les maxima locaux i.e :  
 $\nabla f(x) = 0$  par déplacement de la fenêtre  $Sr, x$

La procédure est la suivante :

1. Choisir le rayon  $r$  de la fenêtre de recherche.
2. Choisissez l'emplacement initial de la fenêtre.
3. Calculez le vecteur de « shift mean » et translatez la fenêtre de recherche par cette quantité (somme).
4. Répétez-vous jusqu'à la convergence.

La convergence est garantie dans [Com99].

Lorsque cette procédure est appliquée à tous les points, les points de convergence s'agrègent en grappe qu'il faut fusionner.

Ceux-ci sont les modes détectés (maxima) et leurs points associés définissent le bassin d'attraction. Les groupes sont délimités par les frontières des bassins et peuvent ainsi avoir des formes arbitraires.

L'algorithme de segmentation proposé ici utilise un espace de 5 caractéristiques : Espace des couleurs  $L*u*v$  ( au lieu de RGB car plus commode pour calculer les distances dans l'hypersphère qui constitue la fenêtre de « shift Mean »)

Les 2 autres sont les coordonnées du pixel dans l'image. On procède en deux étapes :

- 1- l'image est filtrée par « Shift Mean », qui préserve la discontinuité
- 2- les bassins d'attraction localisés dans  $r/2$  sont fusionnés récursivement jusqu'à convergence, les bassins les plus larges qui en résultent sont déclinés en région dont la valeur des pixels est celle de la moyenne (centre).

On peut avoir plus de détails dans [Com99] et [Com02]

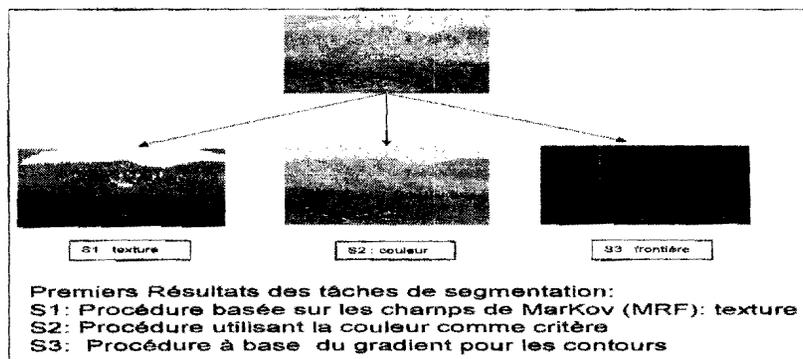
- **Segmentation par les textures** basée sur Champs de Markov cachés (MRF). Nous avons utilisé ici une version proposée par S. DERRODE dans le Projet SIR et présentée récemment dans [Der02]. L'algorithme peut être résumé en 3 phases:

**Phase1** Initialisation: lecture d'une image au format **pgm** (8 bits), calcul de l'histogramme, segmentation initiale par l'algorithme K-means et première estimation des paramètres  $\theta=(\alpha,\beta)$  ou  $\alpha$  désigne les paramètres de régularité de la modélisation markovienne et  $\beta$  désigne les paramètres du mélange (les moments d'ordre 1 à 4 dans le cas du système de Pearson).

**Phase2** Estimation des paramètres: La phase d'estimation des paramètres est réalisée par l'algorithme « Iterative Conditional Estimation » (ICE)[Pie94]. Chaque itération  $n$  fournit une estimation  $\theta_n$  des paramètres  $\theta$

**Phase3** Segmentation: Cette phase permet de segmenter l'image originale à partir des paramètres estimés  $\theta_N$  (à la sortie de ICE) dans un contexte bayésien. L'algorithme utilisé est le « Maximum Posteriori Mode » (MPM).

La segmentation de l'image selon les trois critères : couleur, texture et frontières donne les zones de la figure 6 :



**Figure 6**

Après reparamétrage des procédures afin de ne garder que les zones de grandes tailles et les contours les plus longs (on a fixé un seuil pour la taille et un seuil pour la longueur), on a obtenu comme résultats pour les trois cas :

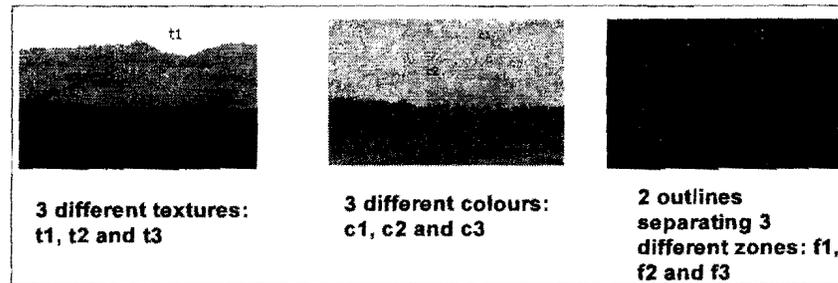


Figure 7

Dans cette segmentation, on remarque que les zones déterminées par les contours ne correspondent pas exactement à celles issues des autres segmentations (textures et couleurs). Par contre ces derniers ont confondu les zones f3 avec t3 et c3 respectivement.

La Combinaison des critères dans le processus de segmentation (voir architecture) permet de corriger ce défaut et de déterminer les quatre zones :  $z1 = t1-c1-f1$ ,  $z2 = t2-c2-f2$ ,  $z3 = t3-c3-f2$  et  $z4 = t3-c3-f3$

Le résultat est le suivant :



Figure 8

#### 4.2. Tâche de la Catégorisation :

Dans cette phase le système doit fournir les objets prototypes correspondant aux zones. A ce niveau, les seules caractéristiques permettant de sélectionner la liste des prototypes sont la texture, la couleur et la forme.

Prenons par exemple la zone **z3** (après son extraction de l'image d'origine) :

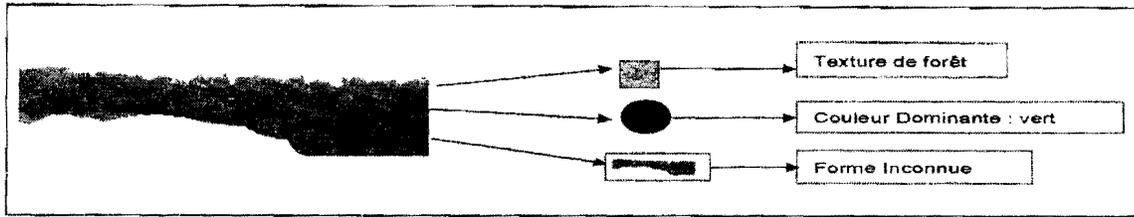


Figure 9

Le processus de la catégorisation pour cette zone se déroule selon le schéma suivant :

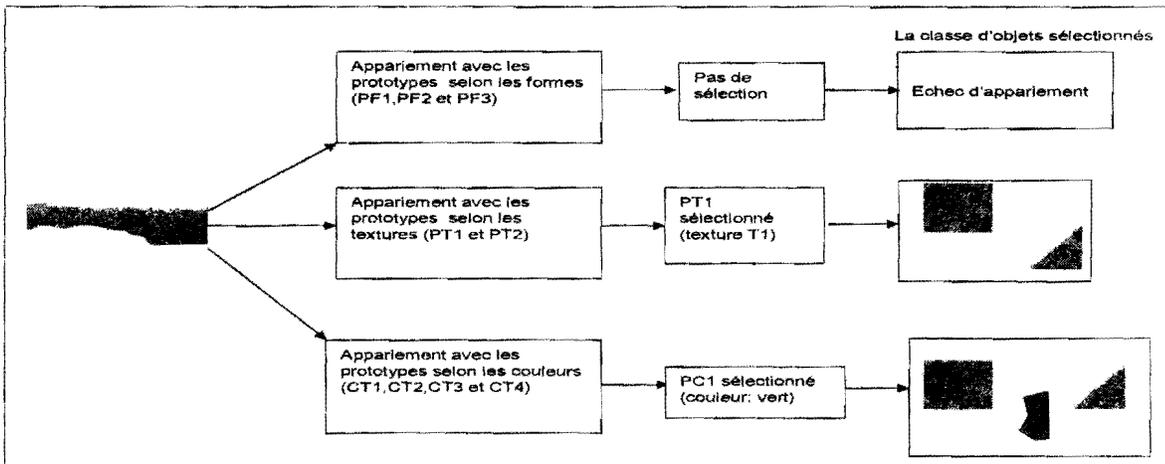


Figure 10

#### 4.3. Phase de décision :

Une simple intersection des 2 classes sélectionnées donnera deux prototypes P1 et P2 ce qui rend plus facile et plus rapide la partie de reconnaissance (deux objets).

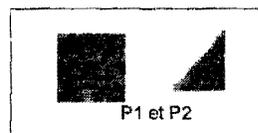


Figure 11

Le processus continue de la même façon pour les autres zones dont nous résumons les résultats:

Dans cet exemple, simplifié pour illustrer notre pensée, l'image à l'entrée est constituée de quatre zones identifiables par les couleurs et les textures:

Les couleurs : Marron (M), Bleu (G) et Vert (V).

Les textures (approximatives) :

T1 (primitif pour les forêts)

T2 (primitif de texture pour les terrains nus)

T3 (primitive de texture pour les végétations)

La zone végétation est n'a pas été déterminée par la segmentation des textures, mais retrouvée par combinaison. Par contre les frontières entre forêt et terrain nu (montagne) n'ont pas été décelées par les contours (plusieurs chaînes de contours trop petites éliminées dans la phase de fusion).

Cet exemple est un des tests pour les outils de base que nous avons développés en utilisant les critères simples (couleur, texture, contours).

<p><b>Résultats :</b> Zone1 : couleur uniforme bleu ( objets sélectionnés : ciel, mer ,...) Zone2 : couleur marron et texture terrain nu (montagne,...) Zone3 : forêt Zone4 : forêt ( mal classée) , forme presque triangulaire</p>
---

## 5. Conclusion

L'introduction de la phase de catégorisation dans un système de vision permet donc de faciliter la tâche à la reconnaissance d'objets. D'habitude, on utilise la forme géométrique (modélisation en trois dimensions) des prototypes dans le traitement d'images. Ici, nous avons tenté d'utiliser d'autres aspects aussi significatifs que la géométrie notamment dans le domaine d'imagerie aérienne. Ce domaine couvre des applications diverses telles que la cartographie, l'agriculture, l'océanographie etc., et l'automatisation du traitement d'images constitue l'outil incontournable des praticiens. Un système complet de vision doit avoir une représentation interne très souple permettant la prise en compte des caractéristiques variées et les plus informatives. C'est à ce que nous voulons aboutir. Nous avons mis en avant les travaux intéressants en catégorisation qui suggèrent la segmentation des catégories en niveaux. De notre part, nous avons proposé une architecture incluant les niveaux. L'architecture basée sur le modèle préconisé nous a permis d'organiser les tâches de la segmentation et de la sélection de prototypes dans cette phase de la catégorisation. La compétition entre plusieurs prototypes permet de livrer rapidement ceux qui sont considérés pertinents à la phase de la reconnaissance. La part du parallélisme n'est pas négligeable, ce qui réduit considérablement le temps de réponse. Nous compléterons (dans un prochain papier) par l'introduction d'un mécanisme permettant d'organiser les tâches en agents intelligents pour bénéficier d'autres propriétés intéressantes comme la coopération et la compétitivité.

## 6. Références :

- [Air95] Airault S. & Jamet O, « Détection et restitution automatique du réseau routier sur images aériennes » *Traitement du Signal*, Vol. 12 Numéro 2 1995.
- [Art02] M. Artaj, M. Jogan & A. Leonardis. *Incremental PCA for On-line Visual Learning and Recognition*. In Proc. of Int. conf. on Pattern Recognition (ICPR'02), 2002.
- [Aya86] N. Ayache and O.D. Faugeras, HYPER: A new approach for the recognition and positioning of two-dimensional objects, *IEEE Trans, PAMI*, 8(1) January 1986, p 44-54.
- [Bai97] C. Baillard « Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains. Détection et caractérisation du sursol », Thèse de Doctorat (Spécialité : Signal et Images), l'Ecole Nationale Supérieure des Télécommunications, France, 1997.
- [Bas96] Basri R. « Recognition by Prototypes » *International Journal of computer vision*, 19(147-168). (1996)
- [Ber96] M. Berthod, Z. Kato, S. Yu et J. Zerubia. Bayesian image classification using Markov random fields. *Image and Vision Computing*, 14(4) :285–293, 1996.
- [Bie87] Biederman, I. Recognition-by-components: A theory of human image understanding. *Psych. Rev.* 94, 115-147. 1987
- [Bre92] Breuel T. M. « Geometrique Aspects of Visual Object Recognition » PhD Thesis, MIT. (1992)
- [Com99] D. Comaniciu and P. Meer. Mean shift analysis and applications. In *7th International Conference on Computer Vision*, pages 1197–1203, Kerkyra, Greece, September 1999.
- [Com02] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Machine Intell.*, 24:603–619, May 2002.

[Can86] J. F. Canny. "A computational approach to edge detection". *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(6):769--768, November 1986.

[Cha94] Y.L. Chang and X. LI, Adaptive Image Growing, *IEEE Trans. On Image Processing*, 1994, November, 3, 6.

[Che95] J. Chen and A. Kundu, "Unsupervised Texture Segmentation using Multichannel Decomposition and Hidden Markov Models", *IEEE Trans. Image Processing*, vol. 4, no. 5, pp. 603-, May, 1995.

[Cor97] Cord, M., Huet, F., and Philipp, S., "Optimal Adjusting of Edge Detectors to Extract Close Contours", In *Scandinavian Conference on Image Analysis*, Lappeenranta, Finland, 1997.

[Der02] S. Derrode and W. Pieczynski, SAR image segmentation using using generalized Pairwise Markov Chains, *SPIE's International Symposium on Remote Sensing*, September 22-27, Crete, Grece, 2002.

[Der88] R. Deriche. An efficient method to build early image description. In *Proc. International Conference on Pattern Recognition (ICPR'88)*, Rome, Italy, November 1988.

[Des99] X. Descombes, R. Morris, J. Zerubia et M. Berthod. Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood. *IEEE Trans. on Image Processing*, 8(7) :954–963, 1999.

[Ede99] Edelman, S. (1999). *Representation and Recognition in Vision*. MIT Press, Cambridge, MA.

[El K99] O. El kharki, M. Sadgal, A. Ait Ouahman « La Cartographie automatique à partir d'images aériennes (site de Rabat –Salé, Maroc) » première rencontre sur les systèmes électroniques et informatiques SEI'99, organisée par la FST de Fés-Saiss (Maroc) le 13-14 octobre 1999.

[El K00] O. El Kharki, M. Sadgal, A. Ait Ouahman, A. El Himdy, M. Ait Belaid “ Automatic Cartography From Aerial Images (Site of Salé,aMorocco)” International Archive of Photogrammetry And Remote Sensing, Vol. XXXI0II, Part B2, pages 179 – 183, 2000.

[El K02a] El Kharki O. Ait Belaid M. Ait Ouahman A. Sadgal M. Bijaber N. “ L’application de la télédétection et des systèmes d’information géographique à l’élaboration de la carte d’occupation du sol dans la province d’Oujda au Maroc”, Journal Maghrebien de physique, volume 2 N° 1 2002

[El K02b] O. El Kharki, M. Ait Belaid, A. Ait Ouahman, M. Sadgal, N. Bijaber, "Classification of multispectral image using ISODATA algorithm" AMSE Journal, paper N° 01315 (1B), 2002.

[Eri98] Erickson, M.A., & Kruschke, J.K. Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, 127, 107-140. 1998

[Fuk90] K. Fukunaga. *Introduction to statistical pattern recognition* (2nd ed). Academic Press, Boston 1990.

[Gar99] P. Garnesson, G. Giraudon & P. Montesino « Messie : un système multi-spécialiste en vision Application en imagerie aérienne », Rapport de recherche, INRIA, avril 1999.

[Hal98] P. Hall, D. Marshall & R. Martin. *Incremental eigenanalysis for classification*. In Proc. of British Vision Machine Conference (BMVC'98), pages 286-295, 1998.

[Hor95] Hornegger, J. and Niemann, H. “Statistical learning, localization and identification of objects”. In ICCV’95 Fifth International Conference on Computer Vision, pp. 914–919. 1995.

[Hor99] J. Hornegger, D. Paulus, and H. Niemann. Probabilistic modeling in computer vision. In B. Jähne, H. Haussecker, and P. Geissler, editors, *Handbook of Computer Vision and Applications*, volume 2. Academic Press, London, 1999.

[Lop99] A. Lopès, R. Fjørtoft, D. Ducrot, P. Marthon, and C. Lemar'echal. Segmentation of SAR images in homogeneous regions. In C. H. Chen, editor, Information Processing for Remote Sensing. World Scientific Publishing Co., Singapore, 1999.

[Low87] D. G. Lowe, The Viewpoint Consistency Constraint, Int'l J. Comput. Vision, (1):57-72, 1987.

[Mar78] Marr, D. and Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. Proc. R. Soc. Lond. B Biol. Sci. 200, 269--294.

[Mel97] Mel, B., SEEMORE: combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition. Neural Comp. 9, 777-804, 1997

[Mur85] Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. Psychological Review, 92, 289-316.

[Nei67] Ulric Neisser , "Cognitive Psychology", 1967

[Pie94] W. Pieczynski, Champs de Markov cachés et estimation conditionnelle itérative, Traitement du Signal 11 (1994), no. 2, 141-153.

[Pog90] Poggio, T. and Edelman, S. (1990). A network that learns to recognize 3D objects. Nature 343, 263--266.

[Ros76] Rosch E. Mervis C. B, Gray W. D. , Johnson D. M., and Boyes-Braem P. « Basic objects in natural categories », Cognition Psychology, 8 :382-439. (1976).

[Ros78] Rosch E. « Principles of categorization » in Rosch E & Lloyd B editors, Cognition and categorization pp27-48. Erlbaum, Hillsdale NJ. (1978).

- [Sad99] Sadgal M. & EL Fazziki A, « Problèmes d'occupation d'espace : l'approche de l'Intelligence Artificielle », FUR IX, The Ninth International Conference on the Foundation and Applications of Utility Risk an Decision Theory, Université CADI AYYAD MARRAKESH, 5 May 1999.
- [Sad00] Sadgal M. & El Fazziki A. « Rôle De La Catégorisation Dans La Reconnaissance d'objets », MCSEAI'2000, Fès - Morocco, 1-3 November 2000
- [Sad01a] M. Sadgal, A. El Fazziki , “Un Modèle de Catégorisation dans les Systèmes de Vision” TELECOM'2001 & 2èmes JFMMA, Casablanca, 17-18-19 Octobre 2001.
- [Sad01b] Sadgal M. & El Fazziki A. “Une Architecture Multi-Agents Basée Sur la Catégorisation dans le Traitement d'images”, International Conference on Image an Signal Processing, 608-617, Agadir, May 3-5 2001
- [She96] J. Shen. On multi-edge detection. CVGIP, Graphics Models and Image Processing, 58(2):101--114, March 1996.
- [Swa91] Swain M. J. & Ballard D. H. (1991), "Color indexing", Int. J. Comput. Vision, 7:11-32.
- [Tar98a] Tarr, M., Williams, P., Hayward, W., and Gauthier, I. Three-dimensional object recognition is viewpoint-dependent. Nat. eurosci. 1, 275—277, 1998
- [Tar98b] Tarr, M. and Bulthoff, H. (1998). Image-based object recognition in man, monkey and machine. Cognition 67, 1-20.
- [Tup98] F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky. Detection of linear structures in SAR images: Application to road network extraction. IEEE Trans. Geoscience and Remote Sensing, 36(2):434--453, March 1998.
- [Wal97] Wallis, G. and Rolls, E. (1997). A model of invariant object recognition in the visual system. Prog. Neurobiol. 51, 167--194.

[Wal98] Walter V. « Automatic classification of Remote Sensing Data for GIS Database Revision » In IAPRS Vol. 32, Part 4, pp 545-552. (1998).

[Wei95] Weidner U. and Förstner W. « Towards Automatic building Extraction from High Resolution Digital Elevation Models » ISPRS Journal, 50, pp 38-49

[Wid90] B. Widrow and M. A. Lehr; 30 years of adaptive neural networks: Perceptron, madaline and back-propagation, In Processin of the Institute of Electrical and Electronic Engineers, volume 78, pages 1415-1442, 1990.

[Won92] C. S. Won and H. Derin, "Simultaneous Parameter Estimation and Segmentation of Noisy and Textured Images using Markov Random Fields", *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 4, pp. 308-328, 1992.

[Wro87] B. Wrobel and O. Monga. Segmentation d'images naturelles : coopération entre un détecteur contour et un détecteur région. In Actes du Onzième colloque GRETSI, Nice, France, Juin 1987.

#### **LES AUTEURS:**

- Mr Mohammed SADGAL : Professeur Assistant à la Faculté des Sciences Semlalia, Marrakech

Département Informatique, LIARC (Laboratoire d'Intelligence Artificielle et Réseaux des Télécommunications)

- Mr Aziz EL FAZZIKI : Professeur à la Faculté des Sciences Semlalia, Marrakech  
Département Informatique, LIARC (Laboratoire d'Intelligence Artificielle et Réseaux des Télécommunications)

- Mr Abdellah AIT OUAHMAN : Professeur à la Faculté des Sciences Semlalia, Marrakech

Département de Physique, Laboratoire de Traitement de Signal

**Tous les auteurs sont des Enseignants Chercheurs à la faculté des sciences Semlalia, Marrakech**