

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة أبي بكر بلقايد - تلمسان

Université Aboubakr Belkaïd – Tlemcen –

Faculté de TECHNOLOGIE



MEMOIRE

Présenté pour l'obtention du **diplôme** de **MASTER**

En : Télécommunications

Spécialité : Réseaux Mobiles et Services de Télécommunications

Par :

RERBAL Ghouti

FOUKIA Salah Eddine

Sujet

Détection et comptage des véhicules dans des séquences vidéo

Soutenu publiquement, le 12 / 06 /2017, devant le jury composé de :

Mr. HADJILA. Mourad	MCB	Univ. Tlemcen	Président
Mr. BOUABDELLAH. Reda	MAA	Univ. Tlemcen	Encadreur
Mr. MOUSSAOUI. Djilali	MAA	Univ. Tlemcen	Examineur
Melle. DJELTI. Hamida	MCB	Univ. Tlemcen	Examinatrice

Dédicaces

C'est avec une très grande émotion et un immense plaisir que je dédie ce modeste travail :

A mes très chers parents qui m'ont soutenu durant toute la durée de mes études.

Mes frères Yacine, Imad Eddine, Anes et ma sœur Hasna

A tous les membres de ma famille, petits et grands

A tous mes chers amis qui m'ont beaucoup aidé durant ces années d'études

A tous ceux qui m'aiment et que j'aime.

FOUKIA Salah Eddine

Dédicaces

C'est avec une très grande émotion et un immense plaisir que je dédie ce modeste travail :

A mes très chers parents qui m'ont soutenu durant toute la durée de mes études.

*Mes frères Rerbal Samira
Rerbal Remzi*

A tous les membres de ma famille, petits et grands

A tous mes chers amis qui m'ont beaucoup aidé durant ces années d'études

A tous ceux qui m'aiment et que j'aime.

RERBAL Ghouti

Remerciements

Nous tenons tout d'abord à remercier ALLAH le tout puissant et miséricordieux, qui nous a donné la force et la patience d'accomplir ce modeste travail.

*En second lieu, nous tenons à remercier notre encadreur Mr **BOUABDALLAH Réda** pour le sujet qu'il nous a proposé et ses précieux conseils scientifiques durant toute la période du travail.*

*Nos vifs remerciements vont également aux président MR **HADJILA Mourad** et au membre de jury Mr **MOUSSAOUI Djilali** et Melle **DJELTI Hamida** pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions.*

Enfin, nous tenons également à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Résumé :

L'objectif de ce travail présenté dans ce mémoire se situe dans le domaine du traitement d'image et plus particulièrement, la segmentation d'image et détection de contour.

Pour cela nous avons développée une application sous matlab pour compter les véhicules et détecter l'encombrement dans une vidéo surveillance.

Mots clés : *Imagerie satellitaire à haute résolution, traitement des images, reconnaissance d'objets, classification, vidéo de surveillance.*

ملخص:

الهدف من الأعمال التي عرضت في هذه المذكرة يكمن في مجال معالجة الصور، وعلى الأخص، تجزئة الصورة والكشف عن الحافة.

لهذا قمنا بتطوير تطبيق عد للمركبات وكشف الازدحام في فيديو للمراقبة

كلمات البحث: صور الأقمار الصناعية عالية الدقة، معالجة الصور، والتعرف على وجوه، التصنيف، رصد الفيديو.

Abstract:

The purpose of the work presented in this paper is the field of image processing, in particular, image fragmentation and edge detection.

For this we have developed the application of counting of vehicles and detection of congestion in video surveillance

Keywords: *high resolution satellite images, image processing, object recognition, classification, monitoring the video*

TABLE DES MATIÈRES

Dédicaces	i
Remerciements	iii
Résumé	iv
Table des matières	v
Liste des figures et des tableaux	ix
Introduction générale	1
Chapitre I : Suivis Des Objets En Mouvements	2
I.1 Introduction	2
I.2 La géolocalisation	2
2.1 Définition	2
2.2 Techniques de la géolocalisation	2
2.2.1 La géolocalisation par satellite	2
2.2.1.1 Les composants du GPS	3
2.2.1.2 Le Principe de fonctionnement	4
2.2.2 La géolocalisation via la téléphonie mobile	5
2.2.2.1 GPS assisté (A-GPS)	6
2.2.2.2 L'angle d'Arrivée (AOA)	6
2.2.2.3 Le temps d'arrivée (TOA)	7
2.2.2.4 La différence de temps d'arrivée TDOA	7
2.2.2.5 La puissance du Signal	7
2.2.2.6 Identifiant de la cellule (Cell-ID)	8
2.2.3 La géolocalisation par adressage IP	8
2.2.4 La géolocalisation par Wi-Fi	9
2.2.5 La géolocalisation par RFID	9
2.2.5.1 Principe de fonctionnement	10
2.3 Avantages et inconvénients	10
2.4 Services de la géolocalisation	11
2.5 Les domaines d'application dans la géolocalisation	11
2.5.1 Transport de passagers	12
2.5.2 Possibilités fonctionnelles	12
I.3 La télédétection	12
3.1 Définition	12

TABLE DES MATIÈRES

3.2 Les notions fondamentales de la télédétection.....	12
3.3 La télédétection passive et active.....	13
3.3.1 Télédétection passive	13
3.3.2 Télédétection active	14
3.4 Les satellites de télédétection Spot	14
3.5 Les domaines d'applications de la télédétection.....	15
3.5.1 Météorologie et sciences de l'atmosphère	16
3.5.2 Océanographie et études littorales.....	17
3.5.3 Applications terrestres.....	17
I.4 Conclusion	18
Chapitre II : Analyser des images satellitaires	19
II.1 Introduction	19
II.2 Définition d'image.....	19
II.3 Images numériques	19
II.4 Types d'images.....	19
II.5 Catégories d'image numérique	20
II.5.1 Image binaire	20
II.5.2 Image en niveaux de gris	20
II.5.3 Image en couleurs	20
II.5.4 Caractéristiques d'une image numérique.....	21
II.6 Les images satellitaires	23
II.6.1 Les images optiques.....	23
II.6.2 Les images SAR.....	24
II.7 Traitement d'images	24
II.7.1 Acquisition.....	24
II.7.2 Prétraitement des images (Filtrage numérique)	25
II.7.2.1 Filtres linéaires	25
II.7.2.2 Les filtres non linéaires	26
II.8 Segmentation	28
II.8.1 Définition de la segmentation	28
II.8.2 Les méthodes de segmentation	28
II.8.2.1 Segmentation par région.....	29
II.8.2.2 Segmentation par contours	29

TABLE DES MATIÈRES

II.8.2.3 Segmentation par coopération région-contour	29
II.8.2.4 Segmentation par classification.....	29
II.8.3 Domaine d'application de la segmentation	29
II.9 Conclusion	30
Chapitre III : La reconnaissance d'objets dans les images	31
III.1 Introduction	31
III.2 Problème de la reconnaissance d'objets.....	31
III.3 Les principaux domaines d'application.....	31
III.4 La vision d'être vivants	32
III.4.1 Théorie des caractéristiques particulières.....	32
III.4.2 La théorie de la reconnaissance par composants	32
III.5 Les défis majeur de la reconnaissance d'objets	33
III.6 La reconnaissance spécifique versus générique	34
III.7 Le schéma général des méthodes de reconnaissance d'objets	35
III.8 Les traitements d'extraction des descripteurs	36
III.8.1 Les points d'intérêt	36
III.8.1.1 Détection des coins	36
III.8.1.2 Les descripteurs SIFT	37
III.8.1.3 Les blobs	38
III.8.1.4 Régions extrêmes maximales stables (MSER)	39
III.8.2 Les descripteurs contours	39
III.8.2.1 Détection des contours.....	39
III.8.2.2 Extraction des descripteurs contours	40
III.8.2.2.1 Les fragments de contours.....	40
III.8.2.2.2 Les primitives géométriques	41
III.8.3 Les descripteurs	42
III.8.3.1 Les descripteurs de formes	42
III.8.3.1.1 Les descripteurs obtenus par une approche région.....	42
III.8.3.1.2 Les descripteurs obtenus par une approche contours	43
III.8.3.2 Les descripteurs de texture et de couleur.....	44
III.9 Les traitements de haut niveau : reconnaissance d'objets.....	44
III.9.1 L'apprentissage structurel.....	44
III.9.2 L'approche fondée sur l'apparence	45

TABLE DES MATIÈRES

III.9.3 Les modèle par l'apparence seule.....	45
III.10 La reconnaissance d'objets cartographiques sur les images satellites et aériennes	46
III.11 Conclusion.....	47
Chapitre IV : Description de l'application	48
IV.1 introduction	48
IV.2 Description de l'application	48
IV.2.1 Interface principale	48
IV.2.2 Interface secondaire	49
IV.3 Application	49
IV.4 Quelques opérations sur les images pour niveau bas	52
IV.5 Conclusion	52
Conclusion générale	56
Bibliographie	57
Annexe	60

Liste des Figures et des Tableaux

Liste des Figures

Figure I.1 : Schéma de la géolocalisation GPS	3
Figure I.2 : Constellation GPS	3
Figure I.3 : Segment de contrôle	4
Figure I.4 : Exemple de récepteur	4
Figure I.5 : Localisation GPS avec un seul satellite	5
Figure I.6 : Localisation par GPS avec deux satellites	5
Figure I.7 : Localisation par GPS avec trois satellites	5
Figure I.8 : La géolocalisation mobile	6
Figure I.9 : GPS assisté (A-GPS).....	6
Figure I.10 : Localisation par AOA	6
Figure I.11 : Localisation par TOA.....	7
Figure I.12 : TDOA (Time Difference Of Arrival).....	7
Figure I.13 : La puissance du Signal.....	8
Figure I.14 : Identifiant de cellule Cell-ID	8
Figure I.15 : Schéma de notions fondamentales de la télédétection	12
Figure I.16 : Schéma de télédétection passive	13
Figure I.17 : Schéma de télédétection active	14
Figure I.18 : Exemple d'imagerie spot	15
Figure I.19 : Une tempête sur le proche Atlantique le 18 octobre 2002.	16
Figure I.20 : Températures de surface de la mer sur l'Atlantique	17
Figure I.21 : Synthèses décennales de l'indice de végétation	18
Figure II.1 : Image binaire.	20
Figure II.2 : Image en niveaux de gris	20
Figure II.3 : Le voisinage (3*3) d'un pixel.....	21

Liste des Figures et des Tableaux

Figure II.4 : Exemple des histogrammes de 3 images différentes.....	22
Figure II.5 : Description de notre démarche pour la détection de bateaux.....	23
Figure II.6 : Exemple d'image SPOT 5 panchromatique de résolution 5m.....	24
Figure II.7 : Les différentes techniques de filtrage.....	25
Figure II.8 : Éléments structurants.....	26
Figure II.9 : Opération d'ouverture (érosion, dilatation).....	27
Figure II.10 : Opération de fermeture (dilatation, érosion).	27
Figure II.11 : Filtre morphologique (érosion, dilatation).	28
Figure II.12 : Schéma représentant les techniques de segmentation	28
Figure III.1 : (a). Exemples de géons, (b). Quelques objets composés de géons.....	33
Figure III.2 : Le schéma général des méthodes de reconnaissance d'objets.....	35
Figure III.3 : Les coins détectés sur une image satellitaire	37
Figure III.4 : Les points d'intérêt et les descripteurs locaux.....	37
Figure III.5 : SIFT, calcul de la signature d'un point d'intérêt.....	38
Figure III.6 : Les fragments de contours de quelques objets courants.....	39
Figure III.7 : Les 35 kAS les plus fréquents du dictionnaire de codes.....	41
Figure III.8 : Les segments linéaires par la détection des événements significatif.....	42
Figure III.9 : Les images d'apprentissage artificielles créées par une grammaire.....	45
Figure III.10 : La conception d'objet dans les approches basées sur l'apparence.....	45
Figure III.11 : Les descripteurs radiaux et angulaires.....	46
Figure IV.1 : Interface de reconnaissance d'objets et détection des véhicules.....	48
Figure IV.2: Interface de traitement d'image en niveau bas.....	49
Figure IV.3 : Image prise de la vidéo.....	50
Figure IV.4: Segmentation.....	50
Figure IV.5 : Éliminé les pixels isolés	50

Liste des Figures et des Tableaux

Figure IV.6 : Détection de contour.....	51
Figure IV.7: Détection des véhicules.....	51
Figure IV.8: Chargement d'image de vidéo.....	52
Figure IV.9: a) Niveau de gris, b) l'histogramm.....	52
Figure IV.10 : a) image améliorer, b) dilatation d'histogramme.....	53
Figure IV.11: Filtre passe-bas.....	53
Figure IV.12: Filtre passe-haut.....	53
Figure IV.13: Filtre passe-band.....	54
Figure IV.14: Filtre morphologie.....	54
Figure IV.15: Filtre dilatation.....	54
Figure IV.16: Filtre érosion.....	55
Figure IV.17: Filtre médian.....	55

Liste des Figures et des Tableaux

Liste des Tableaux

Tableau I.1 : Avantages et inconvénients de chaque technique.	11
Tableau I.2 : Les bandes de l'image SPOT multi spectrale	14
Tableau I.3 : Différentes Applications de la télédétection.....	16

Introduction générale

Avec la généralisation de l'utilisation des images numériques, l'analyse du mouvement dans les images s'est révélée être un outil indispensable pour des applications aussi diverses que la vidéo surveillance, la compression vidéo, l'imagerie médicale, la robotique, l'interaction homme- machine, l'analyse de séquences sportives...etc. En effet, les zones de mouvement d'une séquence d'images correspondent souvent à des événements sur lesquels un système de vision doit se focaliser. L'analyse du mouvement est un vaste sujet qui englobe un certain nombre de problématiques.

Nous pouvons notamment citer :

- la détection des objets en mouvement, c'est-à-dire la détection d'un ensemble de régions d'intérêt en mouvement dans la scène tridimensionnelle observée.
- l'estimation du mouvement, qui consiste à estimer, à partir d'une séquence d'images, le mouvement apparent des objets composants une scène tridimensionnelle.
- le suivi de primitives ou de régions, dont le but est de déterminer la position de chaque primitive ou région dans l'image à chaque instant.
- la reconnaissance et la modélisation d'activités ou de gestes. Les trois premières problématiques (détection du mouvement, détection des objets en mouvement et segmentation basée mouvement), sont en général une première étape pour des outils automatiques de vision par ordinateur. Ces outils peuvent avoir pour vocation, soit uniquement de détecter, soit de détecter et reconnaître, soit de détecter et suivre des objets pour analyser le comportement ou la trajectoire de ces objets.

Le but de ce travail est la détection et le suivi d'objets en mouvement dans une séquence vidéo. L'objet en mouvement est imaginé comme ayant un état interne propre à chaque trame dans la séquence. Nous décomposons ce travail en quatre chapitres.

- ✓ Le premier chapitre on présente des méthodes de détection et de suivis d'objets en mouvement. Nous différencions les techniques de géolocalisation et télédétection.
- ✓ Dans le deuxième chapitre, on présente des généralités sur l'image numérique et les techniques d'analyse en traitant une image.
- ✓ Le troisième chapitre a pour objet de présenter d'abord les approches et les méthodes principales de reconnaissance d'objets en général. Nous avons décrit les différents principaux domaines d'application et les défis majeurs de la reconnaissance d'objets dans les images satellitaire.
- ✓ Le quatrième chapitre a pour l'objet de traiter une séquence vidéo pour détecter des véhicules et reconnaissance d'objet au niveau haut. Ensuite traiter des images prises de la vidéo au niveau bas en appliquant les différentes méthodes de traitement d'image

CHAPITRE I :
Suivis Des Objets
En Mouvements

I.1 Introduction

Ce chapitre a pour but de donner une vision générale sur la géolocalisation et la télédétection, qui sont des domaines dans lesquels nous essayons de détecter la position et suivre des objets avec précision.

La géolocalisation joue un rôle très important dans notre société, puisque la plupart des activités et décisions humaines ont une dimension géographique.

Ainsi, la télédétection de nos jours ouvre de nouvelles techniques permettant d'identifier et de suivre les caractéristiques du globe terrestre. Dans ce chapitre nous allons essayer très brièvement de la définir.

I.2 La géolocalisation

1.2 Définition

A l'origine, La géolocalisation est une technologie avancée qui permet de collecter des informations permettant de localiser un objet ou une personne sur une carte, à l'aide de positions géographiques.

Cette technologie a vu le jour en Amérique, en 1993, elle a été Créé pour les besoins de l'armée américaine, depuis quelques années, cette technologie concept est de plus en plus utilisé pour des applications de gestion de flottes et de sécurisation de véhicules.

2.2 Techniques de la géolocalisation

La géolocalisation est contient plusieurs techniques différentes, qui touchant aux domaines d'activités complémentaires. Dans la partie qui suit, on va regarder chacune des techniques de géolocalisation les plus importantes.

2.2.1 La géolocalisation par satellite

Actuellement le système le plus connu et le plus utilisé est le système américain GPS (Global Positioning System). Nous allons donc expliquer brièvement le fonctionnement global de ce système, qui est le modèle de la géolocalisation par satellite.

Le système GPS est composé de trois parties distinctes : les satellites en orbites autour de la terre, des stations de contrôle et les récepteurs à la disposition des utilisateurs [1].

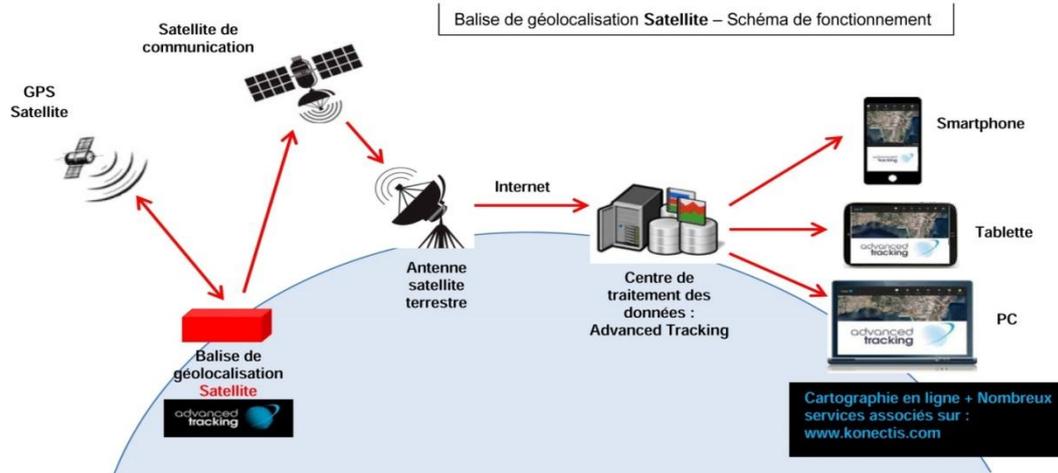


Figure I.1 : Schéma de la géolocalisation GPS [2].

2.2.1.1 Les composants du GPS

Le système GPS comporte trois segments :

- **Segment spatial**

Il est constitué actuellement d'une constellation de 31 satellites (NAVSTAR pour Navigation Satellite Timing And Ranging). Ces satellites évoluent sur 6 plans Orbitaux ayant une inclinaison d'environ 55° sur l'équateur. Ils suivent une orbite quasi circulaire à une altitude de 20 000 à 20 500 km qu'ils parcourent en 11 heures 58 minutes 02 secondes, soit un demi jour sidéral. Ainsi les satellites, vus du sol, reprennent la même position dans le ciel au bout d'un jour sidéral.

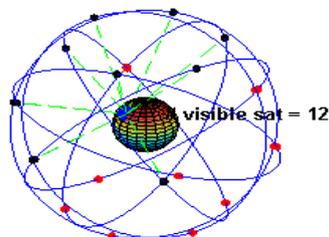


Figure I.2 : Constellation GPS

Il existe plusieurs catégories de satellites GPS:

Bloc I: Satellites lancés entre 1978 et 1985. Tous les satellites du Bloc I sont maintenant hors service excepté un seul, qui est activé de manière périodique. Leur Durée de vie est de 4,5 ans. La principale différence entre ces satellites et les générations suivantes est l'impossibilité de dégrader volontairement le signal transmis.

Bloc II et IIa: Seconde génération de satellites GPS mis sur orbite à partir de 1985. Ils ont la capacité de dégrader le signal émis et leur durée de vie est de 7,5 ans.

Bloc IIr : Ils ont été construits pour avoir une durée de vie de 10 ans. Ils sont capables de communiquer entre eux et ont été mis sur orbite depuis 1996 afin de maintenir une constellation complète.

Bloc IIF : Les satellites Bloc IIF (Follow-On) construits par Boeing seront lancés à partir de 2004. [2]

- **Segment de contrôle :**

Le segment de contrôle comprend cinq stations de poursuite situées à Colorado Springs, Hawaii, Ascension Island, Diego Garcia et Kwajalein (figure ci-dessous). Ces stations sont les yeux et les oreilles du système GPS, elles vérifient l'état des Satellites lors de leur passage au dessus d'elles. Ces stations transmettent ensuite leurs données à la station principale de Colorado Springs. C'est là que les paramètres décrivant l'orbite des satellites et la qualité des horloges embarquées sont estimés, la vérification de la santé des satellites et la détermination d'un repositionnement éventuel sont également contrôlés. Cette information est alors renvoyée à trois stations de chargement qui transmettent l'information aux satellites. Grâce à la répartition uniforme des stations de contrôle, tous les satellites GPS sont captés à 92% du temps [3].

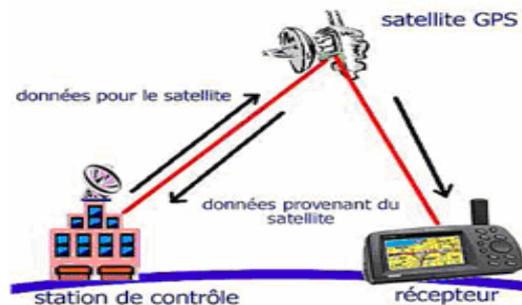


Figure I.3: Segment de contrôle

- **Segment utilisateur**

Le segment utilisateur est constitué de récepteurs qui ont été conçus afin de décoder le signal transmis par les satellites pour déterminer la position, la vitesse et Le temps de l'utilisateur [3].



Figure I.4 : Exemple de récepteur

2.2.1.2 Le Principe de fonctionnement

Le principe du positionnement GPS est très proche du principe de triangulation. La vitesse de transmission des signaux émis par les satellites est égale à celle de la lumière. Chaque signal intègre une éphéméride avec son heure de départ. On mesure donc la distance entre l'utilisateur et un certain nombre de satellites de positions connues grâce au temps qu'a mis chaque signal à parvenir Jusqu'à votre GPS. On définit ainsi des sphères centrées sur des satellites et dont L'intersection donne la position. Le récepteur GPS est capable d'identifier le satellite qu'il utilise à l'aide du signal pseudo aléatoire émis par chaque satellite. Il Charge, à l'aide de ce signal, les informations sur l'orbite et la position du satellite. Pour mesurer la distance qui sépare le satellite du GPS, on mesure le temps T mis par le signal pour aller de l'un vers l'autre. Les schémas ci-dessous montre le calcul de la distance du satellite au récepteur GPS dans le cas d'un satellite, deux satellite et trois satellites [2].



Figure I.8 : la géolocalisation mobile

2.2.2.1 GPS assisté (A-GPS)

La station mobile mesure le temps d'arrivée des signaux envoyés par trois satellites ou plus, les données reçues par le récepteur sont ensuite envoyées via connexion radio aux BTS, ceci tend à résoudre les problèmes liés à la consommation d'énergie abordés en section, puisque le traitement du signal ne se fait pas directement par la station mobile [3].

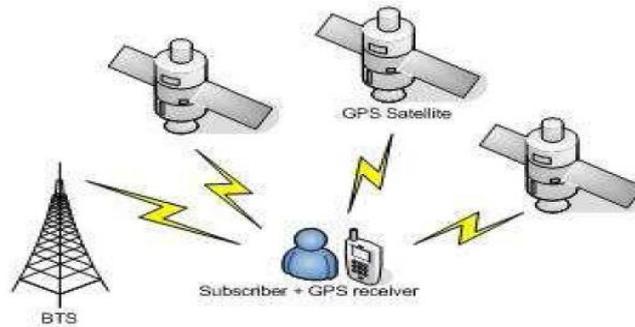


Figure I.9 : GPS assisté (A-GPS)

2.2.2.2 L'angle d'Arrivée (AOA)

En mesurant les angles d'arrivée des signaux transmises par au moins deux BTS, il est possible de déterminer la position de la station mobile par triangulation. La majeure contrainte de cette méthode est que celle-ci requière un environnement libre d'obstacle, cette méthode est donc peu efficace en zone urbaine [3].

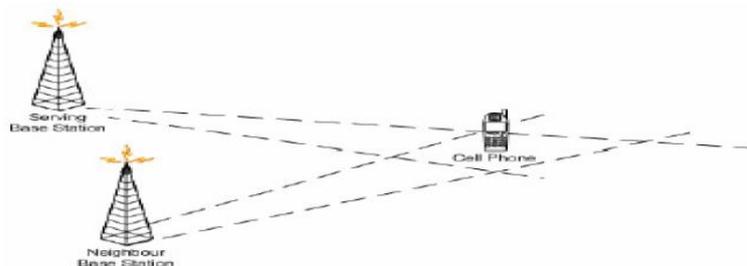


Figure I.10 : Localisation par AOA

2.2.2.3 Le temps d'arrivée (TOA)

TOA calcule le temps de transmission entre la station mobile et le BTS et vice versa. Considérant le fait que le temps de propagation d'une onde radio est connu, il est alors possible d'estimer la distance qui sépare la station mobile du BTS. Cette méthode permet de localiser l'utilisateur dans un cercle qui a pour rayon la distance qui sépare le BTS de la MS et qui a pour centre le BTS [3].

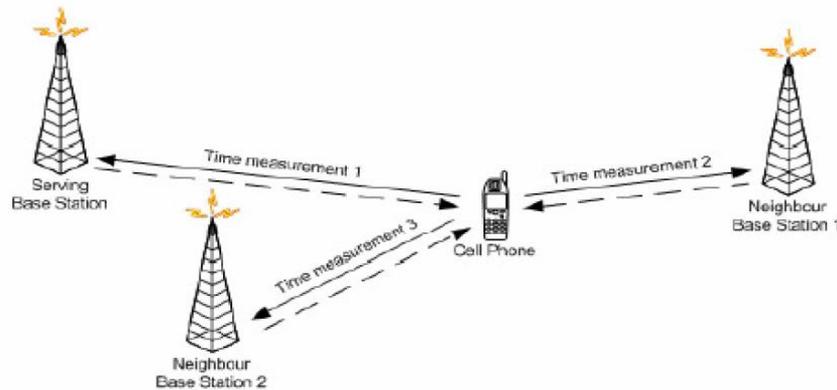


Figure I.11 : Localisation par TOA

2.2.2.4 La différence de temps d'arrivée TDOA (Time Difference Of Arrival)

TDOA mesure le temps de transmission entre la station mobile et trois BTS simultanément, la station mobile se situe à l'intersection des trois cercles. C'est cette méthode qu'on va utiliser par la suite pour étudier le positionnement dans GSM et CDMA [3].

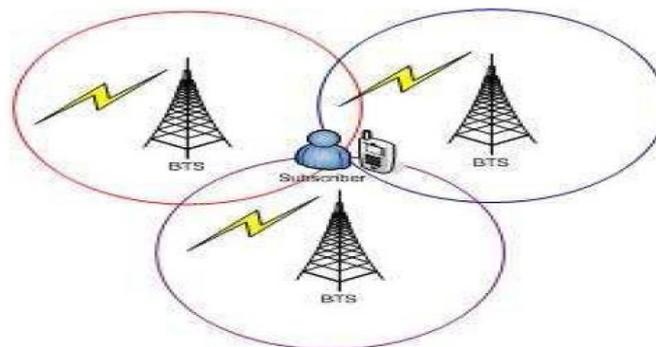


Figure I.12: TDOA (Time Difference Of Arrival)

2.2.2.5 La puissance du Signal

Cette méthode utilise la force du signal pour déterminer la position de la station mobile. Si la puissance de transmission de la BTS et que la puissance de réception de la station mobile sont connus, il est possible d'estimer la distance qui sépare ces deux équipements. Encore une fois la station mobile se situera sur un cercle qui aura pour rayon la distance entre la BTS et la station mobile et pour centre la BTS. Si la puissance du signal est estimée entre trois BTS, la station mobile se situera à l'intersection des trois cercles [3].

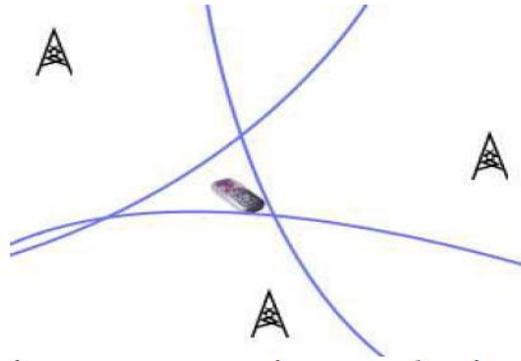


Figure I.13 : La puissance du Signal

2.2.2.6 Identifiant de la cellule (Cell-ID)

Cette méthode est basée sur le réseau GSM et plus particulièrement sur l'identification de la cellule. Elle utilise l'identifiant de la BTS dans lequel la station Mobile est actuellement enregistrée. La topologie du réseau GSM est ensuite utilisée afin d'estimer la position du mobile. La précision de cette méthode dépend principalement de la taille de la cellule ainsi que sur l'environnement avoisinant (C'est-à-dire rural ou urbain) [3].

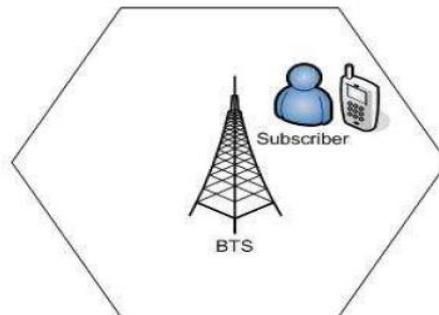


Figure I.14 : Identifiant de cellule Cell-ID

2.2.3 La géolocalisation par adressage IP

La Géolocalisation par IP et implicitement à l'aide des réseaux d'internet par Wifi découle de la manière d'implémentation de l'Internet à l'heure actuelle. Pour comprendre le fonctionnement, il faut tout d'abord savoir comment on a fait jusqu'à présent pour répartir les adresses IP.

Au niveau global, la répartition des adresses IP est faite par l'ICANN, « Internet Corporation for Assigned Names and Numbers » (en français, la Société pour l'attribution des noms de domaine et des numéros sur Internet). Chaque pays se voit attribué une certaine tranche des adresses IP, pouvant être ensuite utilisé par les fournisseurs d'accès à l'Internet du pays respectif. Ensuite, à l'aide d'une base des données, on peut savoir la « nationalité » d'une adresse IP.

En outre, l'ICANN se charge de réguler les domaines de premier niveau, c'est-à-dire un domaine qui est au sommet de la hiérarchie des domaines (par exemple « .fr », « .org » etc.).

On a vu comment on faisait au niveau global pour localiser au moins le pays d'origine d'un ordinateur possédant telle ou telle adresse IP. Maintenant, en chaque pays il existe une agence chargée de réguler l'accès à l'Internet. En France par exemple on retrouve l'ARCEP « L'Autorité de régulation des communications électroniques et des postes ».

En ce qui concerne l'adressage IP, l'ARCEP alloue à chaque fournisseur d'accès à l'Internet un tronçon des adresses utilisable en France. C'est ici que les choses deviennent moins claires. En principe, on devrait avoir une répartition relativement uniforme des adresses du point de vue géographique, chaque opérateur se chargeant de cataloguer les locations des IP. Et souvent c'est le cas. Mais parfois, par exemple, on se voit attribué un IP de Marseille quand on se trouve à Strasbourg.

Le cas des bornes WiFi publiques est par contre plus facile à gérer. Du moment où on en a nettement moins que des postes fixes avec une certaine adresse IP, l'implémentation et l'organisation de ces bornes étant faites par les autorités locaux, la probabilité d'avoir des bases des données fiables est grande.

Pour résumer, la géolocalisation à l'aide des adresses IP est basé sur le fait qu'on sait, à l'aide des bases des données, où devrait se trouver tel ou tel IP, ses bases étant, à la base gérés par les opérateurs d'Internet [1].

2.2.4 La géolocalisation par Wi-Fi

De la même façon qu'un terminal GSM peut se localiser par la méthode du Cell ID sur un réseau GSM, un terminal Wi-Fi peut utiliser la même méthode en se basant sur les identifiants des bornes Wi-Fi (adresses MAC) qu'il détecte. Il existe des bases de données recensant une multitude de bornes d'accès Wi-Fi ainsi que leur position géographique. Ces bases peuvent appartenir à des entreprises privées ou à des communautés qui les publient gratuitement. Ces bases de données sont construites en utilisant la méthode appelée War Driving, qui consiste à parcourir les rues des villes en voiture avec un ordinateur portable équipé du Wi-Fi et relié à un récepteur GPS, afin de recenser un maximum de points d'accès Wi-Fi [4].

2.2.5 La géolocalisation par RFID

La technologie RFID, en anglais «radio-frequency *Identification* », en français, « *Identification par Radio Fréquence* » permet l'identification des objets, d'en suivre le cheminement et d'en connaître les caractéristiques à distance grâce à une étiquette émettant des ondes radio, attachée ou incorporée à l'objet. A l'aide de cette technologie, on peut réaliser la lecture des étiquettes même sans ligne de vue directe et le passage par de fines couches de matériaux (peinture, neige, etc.). Ainsi, grâce aux « lecteurs RFID », on peut accéder aux données mémorisées sur des marqueurs appelés « étiquettes radiofréquences » (« *RFID tag* » ou « *RFID transponder* » en anglais) [4].

2.2.5.1 Principe de fonctionnement

➤ Les radios-étiquettes :

Les étiquettes RFID sont composées d'une puce, utilisée pour mémoriser les informations, relié à une antenne, qui permet l'accès aux données. En fonction de la manière de conception, on distingue plusieurs types :

- Les étiquettes en lecture seule, non modifiables
- Les étiquettes « écriture une fois, lecture multiple »
- Les étiquettes en « lecture réécriture »

Aussi distingue-t-on des étiquettes « passives », qui n'ont pas besoin d'avoir une source d'alimentation propre, pouvant fonctionner juste à l'aide de l'énergie fourni par la borne-lecteur RFID. Ceci entraîne une contrainte au niveau de la distance maximale à laquelle on peut lire ces étiquettes, qui dans le meilleur des cas peuvent attendre une dizaine des mètres.

Par contre, il existe aussi des « tags actifs et semi-actifs », qui intègrent une batterie. Pour les étiquettes actives la batterie est utilisée pour émettre un signal pouvant être lu à des distances plus grandes, en allant d'une cinquantaine de mètres, jusqu'à presque deux cents mètres. Les étiquettes semi-actives n'utilisent pas leur batterie pour émettre des signaux. Elles agissent comme des étiquettes passives au niveau communication. Mais leur batterie leur permet, par exemple, d'enregistrer des données lors du transport.

Un aspect très important dans les dernières années concerne la miniaturisation des composants, qui entraîne la possibilité de les incorporées dans plus d'objets que c'est possible maintenant. Par contre, il faut prendre en compte que la portée des signaux pour accéder à ces puces diminue avec leur taille, conformément à une loi en carré inverse (une quantité physique est inversement proportionnelle au carré de la distance de l'origine de cette quantité physique).

➤ Les lecteurs RFID

Ce sont des dispositifs actifs, émettant des ondes radio à courte distance, qui servent à activer les étiquettes qui passent devant eux. Ce sont ces ondes qui fournissent l'énergie nécessaire pour accéder à la puce. On distingue plusieurs fréquences utilisées, en fonctions du type d'application visé et des performances attendue.

2.3 Avantages et inconvénients

Il existe plusieurs inconvénients et avantages pour chaque technique de géolocalisation [5]

Les techniques	Avantages	Inconvénients
GPS	Précision à long terme : la précision ne se dégrade pas en fonction du temps, il n'y a pas de dérive	Extérieur uniquement

GSM, UMTS	Tout le monde a un Téléphone	Précision dépendant de la densité des relais GSM Dépendant de l'opérateur
WiFi	Bonne précision Utilise réseau existant Terminaux compatibles	Calibrage Logiciel clients à déployer
RFID	Tags RFID passif peu cher visibilité en temps réel	Déploiement matériel de capteur Position disponible au check points
IP	Gratuit !	Précision ville, nécessite base de données géographique

Tableau I.1 : Avantages et inconvénients de chaque technique.

2.4 Services de la géolocalisation

La géolocalisation peut se révéler utile, voir primordiale, dans de nombreux secteurs d'activités. Les finalités et objectifs fixés sont différents selon ces secteurs, mais grâce à une adaptabilité forte et à de nombreuses options, la géolocalisation répond aux attentes variées d'un grand nombre de professions.

1. Garder le contact.
2. La surveillance.
3. Transport de passagers.
4. Suivi de travailleurs mobiles.

2.5 Les domaines d'application dans la géolocalisation

Les applications de la géolocalisation sont en plein développement, tant au niveau privé qu'au niveau professionnel. De plus, couplées à des systèmes de télérelève intégrés et sur mesure, de vraies applications métier ont rapidement vu le jour [6].

Quelques domaines dans lesquels la géolocalisation est communément utilisée sont listés ci-dessous.

2.5.1 Transport de passagers

- Autobus, Navettes, Bateaux, Taxis

2.5.2 Possibilités fonctionnelles :

- Suivi en temps réel
- Calcul du temps avant passage
- Localisation du véhicule le plus proche ou dans une zone spécifique
- Génération de statistiques (temps d'arrêt moyen, temps de chargement de passagers, calcul du temps de parcours en fonction des plages horaires...)
- Optimisation des horaires et des trajets

I.3 La télédétection

3.1 Définition

La télédétection est l'ensemble des techniques qui permettent, par l'acquisition d'images, d'obtenir de l'information sur la surface de la Terre (y compris l'atmosphère et les océans), sans contact direct avec celle-ci. La télédétection englobe tout le processus qui consiste à capter et enregistrer l'énergie d'un rayonnement électromagnétique émis ou réfléchi, à traiter et analyser l'information qu'il représente, pour ensuite mettre en application cette information.

La période qui commence en 1957 marque les débuts de l'exploration de l'Espace et prépare l'avènement de la télédétection actuelle. Le lancement des premiers satellites, puis de vaisseaux spatiaux habités à bord desquels sont embarqués des caméras, révèle l'intérêt de la télédétection depuis l'espace [7]

3.2 Les notions fondamentales de la télédétection

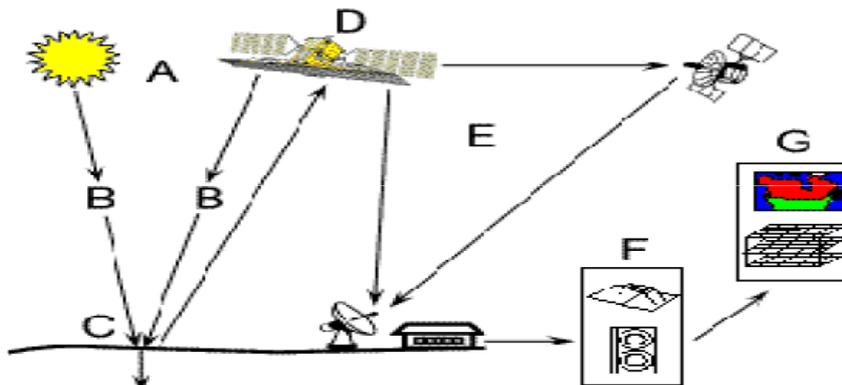


Figure I.15 schéma de notions fondamentales de la télédétection

A. Source d'énergie ou d'illumination

Tout processus de télédétection nécessite une source d'énergie pour illuminer la cible.

B. Rayonnement et atmosphère

Durant son parcours entre la source d'énergie et la cible, le rayonnement interagit avec l'atmosphère. Une seconde interaction se produit lors du trajet entre la cible et le capteur.

C. Interaction avec la cible

Une fois parvenue à la cible, l'énergie interagit avec la surface de celle-ci. La nature de cette interaction dépend des caractéristiques du rayonnement et des propriétés de la surface.

D. Enregistrement de l'énergie par le capteur

Une fois l'énergie diffusée ou émise par la cible, elle doit être captée à distance (par un capteur qui n'est pas en contact avec la cible) pour être enfin enregistrée.

E. Transmission, réception et traitement

L'énergie enregistrée par le capteur est transmise, souvent par des moyens électroniques, à une station de réception où l'information est transformée en images (numériques ou photographiques).

F. Interprétation et analyse

Une interprétation visuelle et/ou numérique de l'image traitée est ensuite nécessaire pour extraire l'information que l'on désire obtenir sur la cible.

G. Application

La dernière étape du processus consiste à utiliser l'information extraite de l'image pour mieux comprendre la cible, pour nous en faire découvrir de nouveaux aspects ou pour aider à résoudre un problème particulier [7].

3.3 La télédétection passive et active

La télédétection comporte deux types :

3.3.1 Télédétection passive

Jusqu'à maintenant, nous avons vu que le Soleil est une source d'énergie pour la télédétection. L'énergie du Soleil est soit réfléchi (la portion visible) ou absorbée et retransmise (infrarouge thermique) par la cible.

Les dispositifs de télédétection qui mesurent l'énergie disponible naturellement sont des capteurs passifs.

Le capteur passif peut seulement percevoir l'énergie réfléchi lorsque le Soleil illumine la Terre.

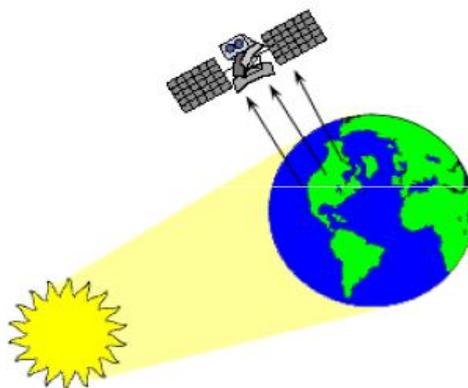


Figure I.16 schéma de télédétection passive

3.3.2 Télédétection active

Un capteur actif produit sa propre énergie pour illuminer la cible : il dégage un rayonnement électromagnétique qui est dirigé vers la cible. Le rayonnement réfléchi par la cible est alors perçu et mesuré par le capteur.

Le capteur actif a l'avantage de pouvoir prendre des mesures à n'importe quel moment de la journée ou de la saison [7].

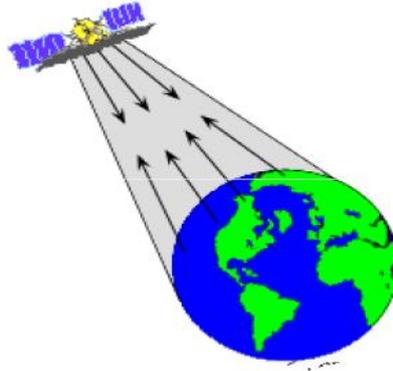


Figure I.17 schéma de télédétection active

3.4 Les satellites de télédétection Spot

SPOT (Système d'observation de la Terre ou Satellite pour l'observation de la Terre) est une famille de satellites de télédétection français civils d'observation de la Terre développés par l'agence spatiale française, le CNES (Centre national d'études spatiales). Cinq premiers exemplaires ont été lancés entre 1985 et 2002. Les images fournies par les satellites sont commercialisées par Astrium. Spot 6 a été lancé en 2012 et SPOT 7 en 2014 [7].

Les caractéristiques de ces satellites définissent la résolution et le mode spectral des images. Il faut savoir que le système SPOT est opérationnel depuis février 1986, et qu'il existe plusieurs générations de satellites avec des caractéristiques différentes.

Le principal avantage des images SPOT est leur disponibilité depuis 1986 avec des capteurs similaires. Cela permet ainsi d'analyser l'évolution de certains phénomènes dans le temps.

De plus, les bandes spectrales des satellites SPOT permettent de suivre l'évolution des couverts végétaux [8].

Une image multi spectrale est une image obtenue a partir de capteurs opérant dans plusieurs bandes spectrales a la fois, tableau suivant :

Bande spectrales	Canal	Longueur d'onde
bande 1	Vert	0.50 -- 0.59 μm
bande 2	Rouge	0.61 -- 0.68 μm
bande 3	proche infrarouge	0.78 -- 0.89 μm

Tableau I.2 : Les bandes de l'image SPOT multi spectrale

Les satellites SPOT 4 et SPOT 5 comprennent une bande supplémentaire bande 4 en moyen infrarouge (1.58 -- 1.75 gm).

Les bandes spectrales les plus utilisées dans les analyses sont le Rouge (bande 2) et l'Infrarouge (bande 3). Cela paraît logique, car ce sont les bandes dans les quelles le contraste entre le sol et la végétation est le plus grand [9].

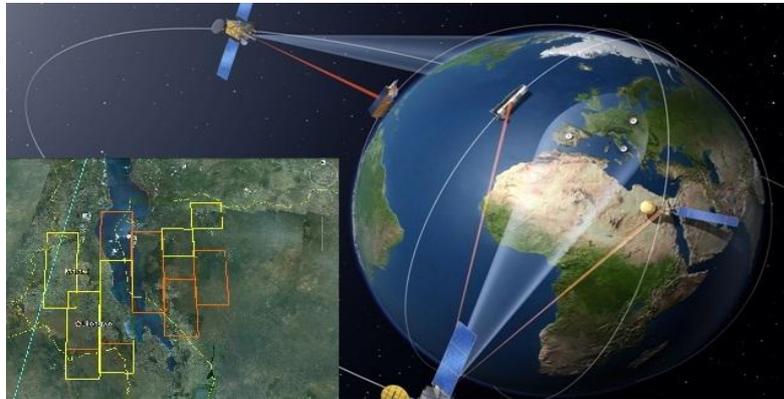


Figure I.18 : exemple d'imagerie spot

3.5 Les domaines d'applications de la télédétection

Le premier grand domaine d'application de la télédétection a été l'étude de l'atmosphère (météorologie et climatologie). L'intérêt de la télédétection dans ce domaine est d'assurer une couverture globale et très fréquemment répétée de la planète entière ; par contre la résolution spatiale n'est pas primordiale pour les applications météorologiques.

Vecteurs	Capteurs	Domaines d'application
Télédétection de l'ATMOSPHERE (Météorologie, Climatologie)		
Satellite géostationnaires (Météosat). Satellite à défilement (NOAA).	Basse et moyenne résolution (on privilégie la répétitivité et la couverture spatiale). Capteurs passifs : visible, infrarouge, microondes. Sondeurs atmosphériques. Dans le futur : radars pluviométriques, lidars (capteurs à laser).	Etude de la nébulosité Mesure des températures Vapeur d'eau et précipitations Eléments du bilan radiatif
OCEANOGRAPHIE Et ETUDES LITTORALES		
Avions, Satellites météorologiques ou de télédétection terrestre. Satellites spécialisés (Nimbus, seasat, ERS-1).	Toutes résolutions selon les espaces considérés (de l'océanographie côtière à l'océanographie globale). Capteurs passifs : visible, infrarouge, microondes. Radars imageurs, radar-altimètre, diffusiomètre.	Analyse de la couleur de l'océan (production biologique, turbidité). Mesures des températures de surface de la mer. Vagues et vents, altitude de la surface (dynamique de l'océan). Glaces de la mer.

APPLICATIONS TERRESRES		
Avion, Satellites à défilement en orbite polaire (Landsat, SPOT).	Surtout haute et très haute résolution spatiale : Photographie aérienne. Capteurs passifs : radiomètres à balayage (domaine optique). Capteurs actif : radars imageurs.	Cartographie régulière et thématique Géologie, prospection minière, géomorphologie. Hydrologie, neige, risques naturels. Agriculture, sylviculture. Urbanisme, aménagement, génie civile.

Tableau I.3 : défèrent Application de la télédétection

3.5.1 Météorologie et sciences de l'atmosphère

- **L'imagerie Météosat**

Les images et animations des satellites météorologiques sur orbite géostationnaire sont les documents de télédétection les plus diffusés auprès du grand public. L'imagerie des satellites européens de la série Météosat en est un exemple.

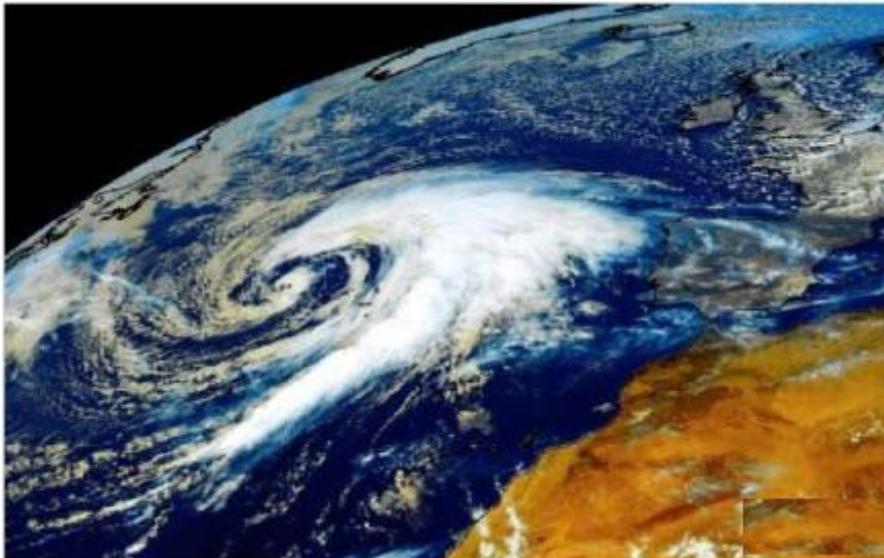


Figure I.19: Une tempête sur le proche Atlantique le 18 octobre 2002. Image Météosat.

Ce type d'imagerie a plusieurs types d'applications :

- Le suivi en temps réel des masses nuageuses et des phénomènes météorologiques (les satellites Météosat fournissent une image toutes les demi-heures.
- L'extraction automatique de paramètres géophysiques qui sont assimilés dans les modèles numériques de prévision météorologique.
- Le traitement en série des images archivées pour l'obtention de séries climatologiques sur la nébulosité et les pluies, la température de surface de la terre ou de la mer, etc.

Des satellites géostationnaires similaires à Météosat couvrent l'ensemble de la planète à l'exception des très hautes latitudes (régions polaires) et permettent une surveillance météorologique globale.

3.5.2 Océanographie et études littorales

L'océanographie est un domaine d'applications de la télédétection remarquable par la diversité des capteurs utilisables et des échelles de travail. L'exemple qui suivent en sont l'illustration.

- **Mesure des températures de surface de la mer**

Les radiomètres infrarouges thermiques à bord de satellites météorologiques sont un outil précis pour la mesure des températures de surface de la mer. Sur une période assez longue, il devient possible de s'affranchir des nuages et de proposer des cartes sur de vastes surfaces. La précision de la mesure est de l'ordre de 0,5°C.

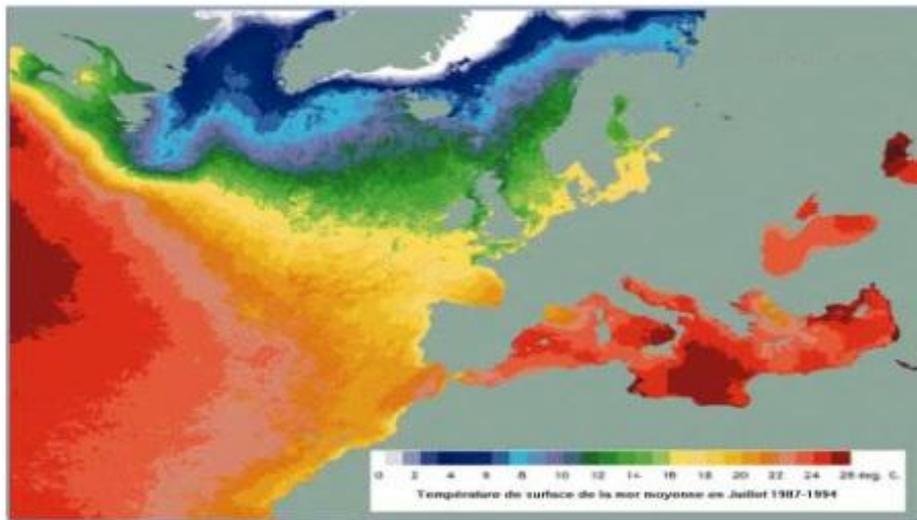


Figure I.20 : Températures de surface de la mer sur l'Atlantique

3.5.3 Applications terrestres

- **Suivi de la végétation à l'échelle planétaire.**

Les applications de la télédétection dans l'étude des surfaces continentales font fréquemment appel à des données à faible résolution spatiale. C'est ici le cas du suivi de la végétation continentale observée par les radiomètres optiques (visible et proche infrarouge) AVHRR des satellites météorologiques de la NOAA. L'activité chlorophyllienne est mesurée à partir d'un indice de végétation [10].

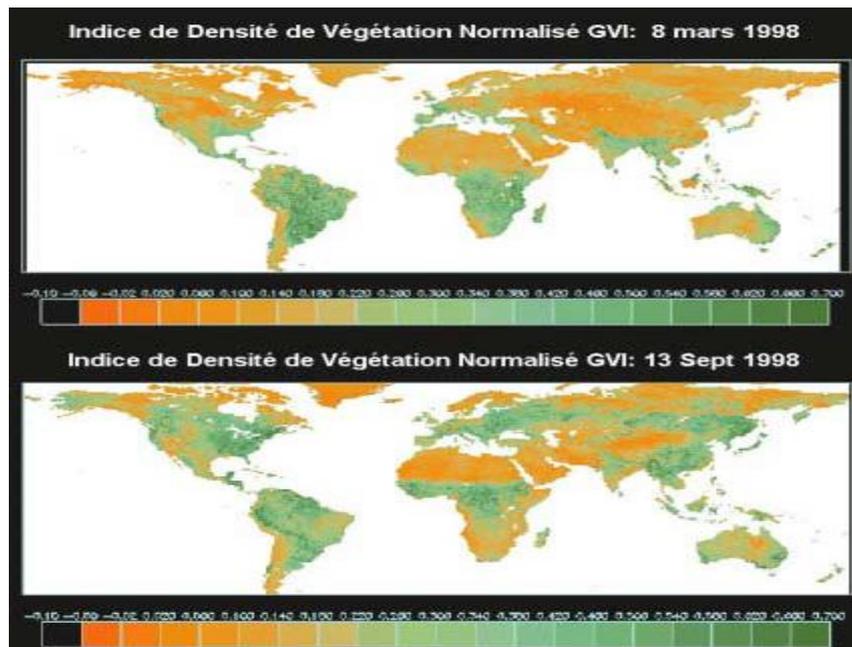


Figure I.21: Synthèses décennales de l'indice de végétation

I.4 Conclusion

Dans ce chapitre, nous avons présenté la géolocalisation et la télédétection avec les différentes technologies et domaines d'applications.

L'objet de chapitre suivant est de faire l'analyse d'une Image satellitaire en utilisant les nouvelles technologies de traitement d'image qui permettent d'améliorer les résultats de l'interprétation des images satellitaires.

CHAPITRE II :

Analyse des

images

II.1 Introduction

Avec la parole, l'image constitue l'un des moyens les plus importants qu'utilise l'homme pour communiquer avec autre. C'est un moyen de communication universel dont la richesse du contenu permet aux êtres humains de tout âge et de toute culture de se comprendre.

L'objectif de l'analyse d'image est de fournir une description quantitative de l'image ou une reconnaissance de forme. On trouve des applications dans de nombreux domaines : sciences des matériaux, sciences de la vie, géologie, robotique...

Le but de ce chapitre, nous allons définir les opérations d'analyser et traiter des images pour extraire des informations et des phénomènes qui permettent à l'être humain de communiquer entre eux.

II.2 Définition d'image

Une image est avant tout un signal 2D (x,y) . Souvent, cette image représente une réalité 3D (x,y,z) .

D'un point de vue mathématique :

- Une image est une matrice de nombres représentant un signal
- Plusieurs outils permettent de manipuler ce signal

D'un point de vue humain :

- Une image contient plusieurs informations sémantiques
- Il faut interpréter le contenu au-delà de la valeur des nombres

II.3 Images numériques

Une image numérique est une image (dessin, icône, photographie, ...) créée, traitée, stockée sous forme binaire (suite de 0 et de 1).

C'est une image dont la surface est divisée en éléments de tailles fixes appelés pixels, ayant chacun comme caractéristique un niveau de gris ou de couleurs prélevé à l'emplacement correspondant dans l'image réelle, ou calculé à partir d'une description interne de la scène à représenter [11].

II.4 types d'images

On distingue généralement deux grandes catégories d'images:

- **Bitmap** (appelées aussi images matricielle) : il s'agit d'images pixellisées, c'est-à-dire un ensemble de points (pixels) contenus dans un tableau, chacun de ces point possédant une ou plusieurs valeurs décrivant sa couleur.
- **Vectérielles** : les images vectorielles sont des représentations d'entités géométriques telles qu'un cercle, un rectangle ou un segment. Ceux-ci sont représentés par des formules mathématiques (un rectangle est défini par deux points, un cercle par un centre et un rayon, une courbe par plusieurs points et une équation).

II.5 catégories d'image numérique

II.5.1 Image binaire

Une image binaire est une image dont les pixels ne peuvent avoir que les valeurs 0 et 1. Le 0 correspond à un pixel noir et 1 à un pixel blanc. Le niveau de gris est donc codé sur un seul bit [12].



Figure II.1 : Image binaire.

II.5.2 Image en niveaux de gris

Dans une image en niveaux de gris, la couleur d'un pixel peut prendre des valeurs allant du noir (0) au blanc (255) en passant par un nombre fini de niveaux intermédiaires obtenus par dégradation du noir. La couleur d'un pixel est ainsi codée sur un octet [13].

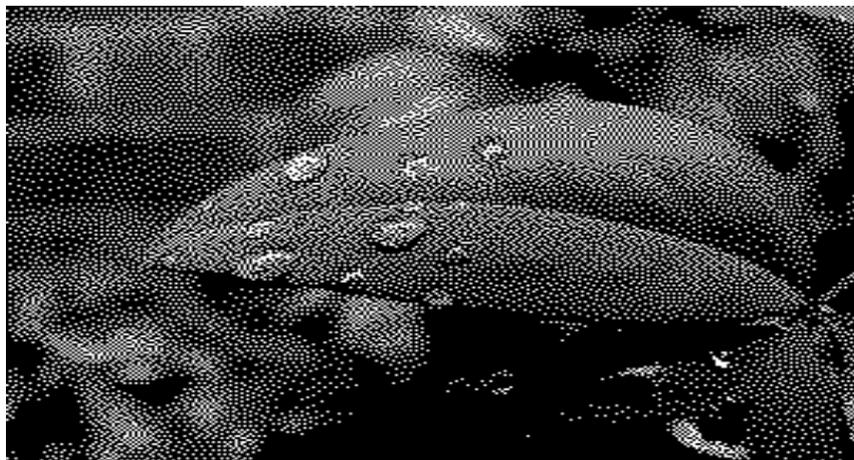


Figure II.2 : image en niveaux de gris

II.5.3 Image en couleurs

Elle est obtenue par la combinaison de trois couleurs dites primaires : rouge, vert et bleu (RVB). Chaque couleur est codée comme une image à niveaux de gris, avec des valeurs allant de 0 à 255. Pour $R=V=B=0$ nous aurons un noir pur et pour $R=V=B=255$ nous aurons un blanc pur. La représentation des images couleurs se fait donc soit par une image dont la valeur du pixel est une combinaison linéaire des valeurs des trois composants couleurs, soit

par trois images distinctes représentant chacune une composante couleur, on distingue généralement deux types d'images : [14]

- Images 24 bits.
- Images à palettes (256 couleurs).

II.5.4 Caractéristiques d'une image numérique

L'image est un ensemble structuré d'informations parmi ses caractéristiques nous pouvons citer les paramètres suivants :

1. Pixel et son voisinage

Une image numérique est constituée d'un ensemble de points appelés pixels. Le pixel représente le plus petit élément constitutif d'une image.

Les pixels du voisinage sont ceux qui entourent le pixel en question (figure suivante). C'est une fenêtre de dimension impaire (3*3, 5*5, ...) [14].

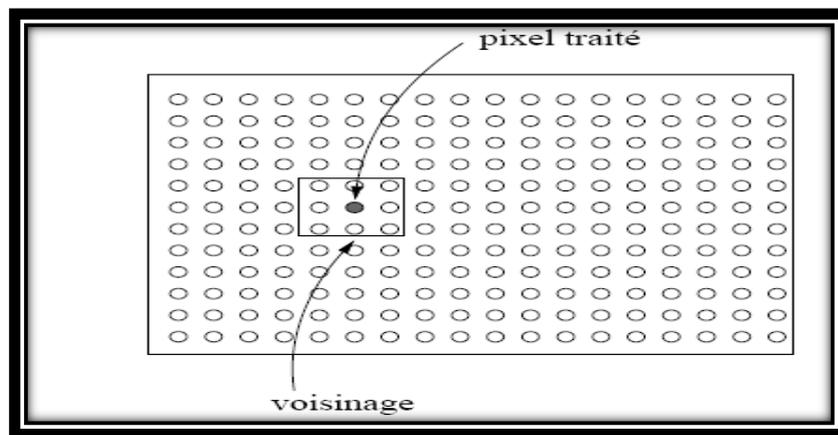


Figure II.3 : le voisinage (3*3) d'un pixel.

2. L'histogramme

L'histogramme d'une image est une représentation graphique en 2D dans laquelle les X représente les différents couleurs (**exemple** (0-255) pour une image en niveau de gris) et le Y représente le nombre de répétition des pixels ayant la même couleur.

Dans cette représentation, un point de coordonnées (x,y) signifie qu'on a X pixels dans tout l'image ayant la même couleur Y [14].

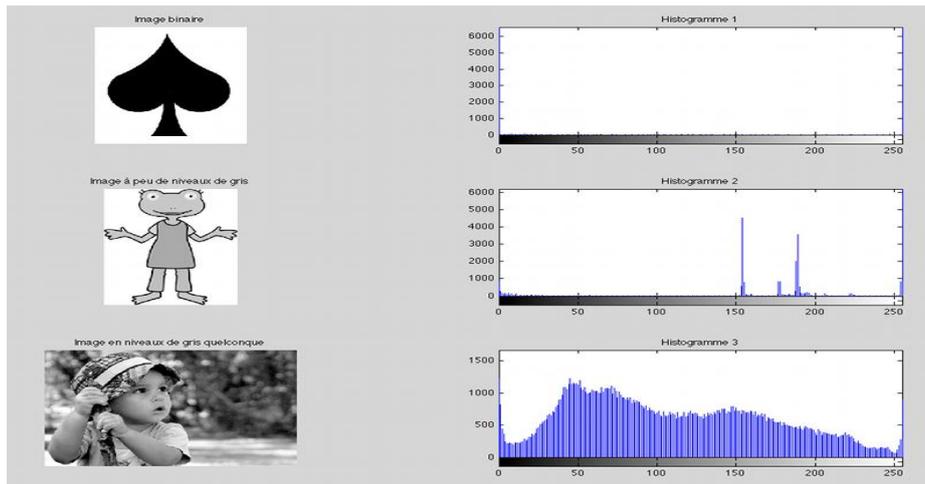


Figure II.4 : Exemple des histogrammes de 3 images différentes.

3. La résolution

Elle est définie par le nombre de pixels par unité de surface (pouce), un pouce représentant 2.54 cm. [14]

4. La luminance

La luminance est le degré de luminosité de chaque point de l'image. Elle est définie comme étant le quotient de l'intensité lumineuse d'une surface [14].

5. Le contraste

C'est l'opposition marquée entre deux régions d'une image, plus précisément entre les régions sombres et claires. Le contraste est défini en fonction des luminances de deux zones d'image [14].

6. Similarité

La similarité est à la base de tout système de comparaison des pixels d'images, il s'agit de définir une fonction de similarité entre les intensités lumineuses des pixels [14].

7. Homogénéité

C'est un concept qui se base souvent sur des critères statistiques (moyenne, variance) [14].

8. Le bruit

Un bruit dans une image est considéré comme un phénomène de brusque variation de l'intensité d'un pixel par rapport à ses voisins, il provient de l'éclairage des dispositifs optiques et électroniques du capteur [15].

9. Le flou

Le flou est la perte de netteté d'une photographie numérique. Ce flou numérique peut venir d'un mouvement soudain du photographe ou bien du capteur de l'appareil numérique [15].

10. La région

C'est un ensemble de pixels connexes et homogènes. Un pixel appartient à une région donnée s'il vérifie les caractéristiques de celle-ci. Une région est limitée par un contour [15].

11. Le contour

Un contour est un ensemble de pixels formant une frontière entre deux ou plusieurs régions voisines, ou la limite entre deux pixels dont le niveau de gris représente une différence significative [15].

II.6 Les images satellitaires

II.6.1 Les images optiques

Ces images sont calculées à partir des données acquises par un capteur passif qui va enregistrer l'énergie solaire réfléchi par la surface terrestre. La nuit, ou en présence de nuages, ce capteur devient aveugle.

On distingue deux type d'images : panchromatique ou multi spectrale (multi bandes).

Dans l'imagerie panchromatique, le capteur mesure la réactance dans une large bande électromagnétique alors que dans l'imagerie multi spectrale, le capteur mesure la réactance dans de nombreuses bandes.

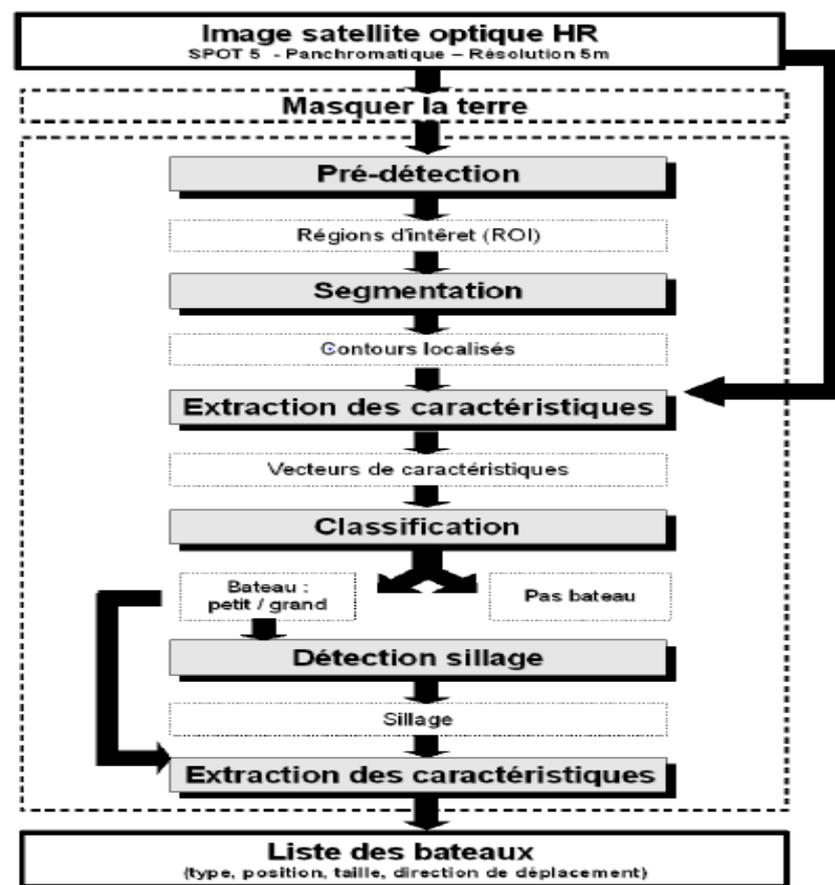


Figure II.5 : Description de notre démarche pour la détection de bateaux.

Ces images présentent un aspect assez semblable à celui d'une photographie et sont donc bien adaptées à une interprétation visuelle [16].

II.6.2 Les images SAR

Ce type d'image est produit par des satellites utilisant un capteur actif qui produit un rayonnement électromagnétique qu'ils vont diriger vers la Terre et enregistrer la part de rayonnement réfléchi par celle-ci. Ce type de capteur réalise des prises de vue de jour comme de nuit et dans toutes les conditions météorologiques, y compris avec une grande couverture nuageuse. Contrairement aux images optiques, l'interprétation de ces images n'est pas simple pour l'œil humain [16].

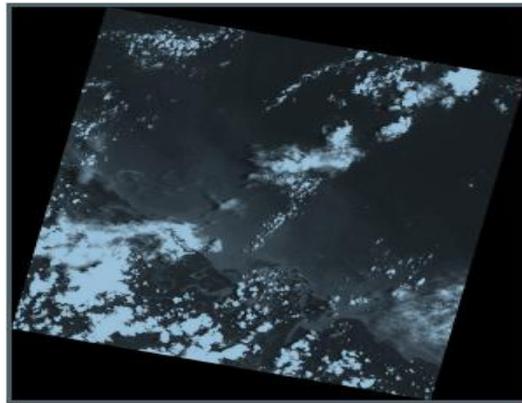


Figure II.6 : Exemple d'image SPOT 5 panchromatique de résolution 5m.

II.7 Traitement d'images

Pour manipuler une image, on travaille sur un tableau d'entiers qui contient les composantes de chaque pixel. Les traitements s'appliquent toujours aux images en niveau gris et parfois aussi sur des images couleur. Nous allons distinguer plusieurs types de traitements (Acquisition, Prétraitement, Segmentation, ...).

II.7.1 Acquisition

L'acquisition d'image constitue un des maillons essentiels de toute chaîne de conception et de production d'images. Pour pouvoir manipuler une image sur un système informatique, il est nécessaire de lui faire subir une transformation qui la rendra lisible et manipulable par ce système. Le passage de cet objet externe (l'image d'origine) à sa représentation interne (dans l'unité de traitement) se fait grâce à une procédure de numérisation.

L'acquisition se fait en plusieurs étapes :

- ✚ Le transfert d'image : c'est le passage du signal optique au signal analogique (électrique).
- ✚ L'échantillonnage du signal analogique.
- ✚ La quantification qui consiste à traduire les échantillons en valeurs binaires utilisant des codes précis [17].

II.7.2 Prétraitement des images (Filtrage numérique)

Le prétraitement est la deuxième étape de notre système de traitement d'image, il a pour but la réduction des bruits induits par les capteurs afin d'améliorer la perception de certains détails. Cette étape renforce aussi la ressemblance entre pixels de même région et la dissemblance des pixels de régions différents surtout dans les cas où le contraste et la luminosité sont faibles. On se propose présenter quelques ces prétraitements.

Les images numériques telles qu'elles sont acquises, sont très souvent inexploitable pour le traitement d'images. Elles contiennent des signaux bruités. Pour remédier à cela, différents prétraitements pour l'amélioration ou la correction sont effectués. Les bruits sont ainsi réduits, les pixels parasites éliminés, et les contours plus définis [17].

Pour résoudre ce problème, on distingue généralement trois types de filtre :

- Filtres linéaires
- Filtres Non-linéaires

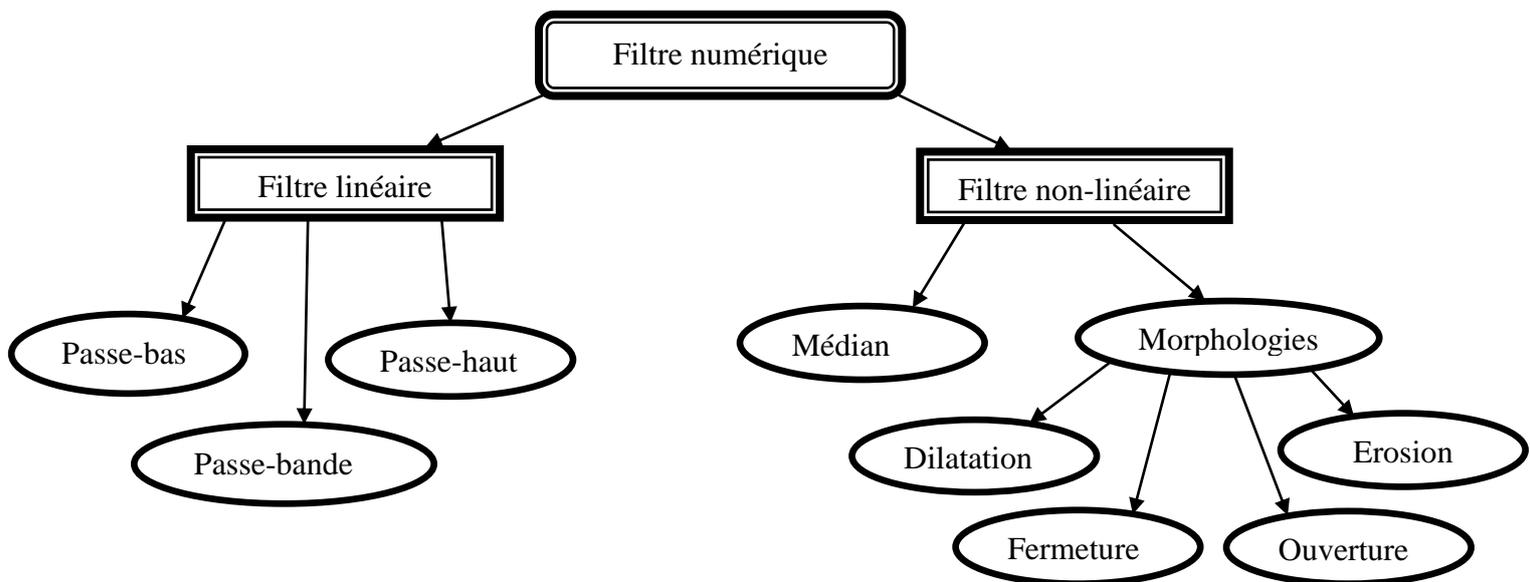


Figure II.7 : Les différentes techniques de filtrage.

II.7.2.1 Filtres linéaires

Les filtres linéaires transforment un ensemble de données d'entrée en un ensemble de données de sortie par une convolution bidimensionnelle qui est une opération mathématique. Ils permettent de supprimer le bruit dans l'image. Chaque filtre a une taille $N \times N$ avec N impair. Parmi les filtres linéaires les plus utilisés on trouve :

a. Les filtres passe-bas :

Dans ce type des filtres, toutes les fréquences qui dépassent une fréquence bien déterminée seront automatiquement éliminées et vont être rejetées. Les autres fréquences (inférieures à la fréquence de seuil) seront acceptées.

b. Les filtres passe-haut

Les filtres passe-haut éliminent tous les fréquences qui au-dessous à la fréquence de seuil et accepte les autres fréquences.

c. Filtre passe-bande (Différentiation)

Ce filtre est un mélange de deux filtres (passe bas, passe haut), donc le filtre éliminent tous les fréquences qui au-dessus a une fréquence max et au-dessous a la fréquence min. Les fréquences entre fréquence max et fréquence min seront acceptées.

II.7.2.2 Les filtres non linéaires

Pour éviter le flou introduit par les filtres passe-bas classiques, une autre famille de méthodes existe, les filtres non linéaires ou filtres d'ordre. Ce sont des filtres dans lesquels la nouvelle valeur du pixel est calculée par une combinaison non linéaire de la fenêtre des pixels avoisinants. Le filtre médian est le filtre non linéaire le plus utilisé :

a. Le filtre médian

Il supprime le bruit dans l'image. Son algorithme est relativement simple, le filtre remplace la valeur de chaque pixel par la valeur médiane du pixel et de ses voisins.

b. Les filtres morphologiques :

Toutes les transformations morphologiques sont définies à l'aide d'un élément structurant et un opérateur. L'élément structurant est un ensemble de pixels qui possèdent les caractéristiques suivantes :

- Une forme géométrique connue (voir la figure suivante).
- Un pixel central noir (voir la figure suivante).
- Un ensemble de pixels voisins au pixel central en blanc (voir la figure suivante).

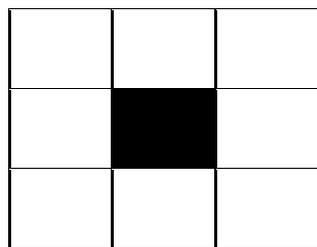


Figure II.8 : Éléments structurant.

Les opérateurs morphologiques les plus usuels sont :

➤ La Dilatation :

Elle permet d'éliminer les points noirs isolés de l'image. Elle élargit les « pics » et comble les « vallées ». La dilatation peut être réitérée jusqu'à ce que l'image ait une valeur constante (la valeur maximale).

➤ **L'Érosion :**

Elle permet d'éliminer les points blancs isolés de l'image. Elle abaisse les «pics » et élargit les «vallées ». L'érosion peut être réitérée jusqu'à ce que l'image ait une valeur constante (la valeur minimale).

➤ **L'ouverture :**

C'est une érosion suivie d'une dilatation, On dit que l'ouverture rase les «pics » l'histogramme sans modifier les « vallées ».

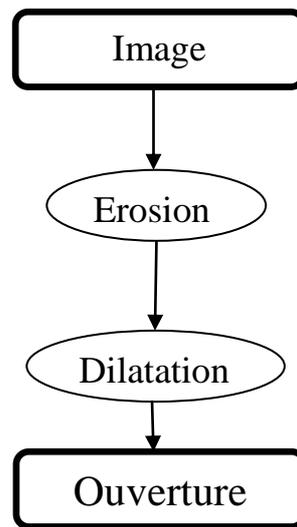


Figure II.9 : Opération d'ouverture (érosion, dilatation).

➤ **La Fermeture :**

C'est une dilatation suivie d'une érosion. On dira que la Fermeture comble les « vallées » sans modifier les «pics ».

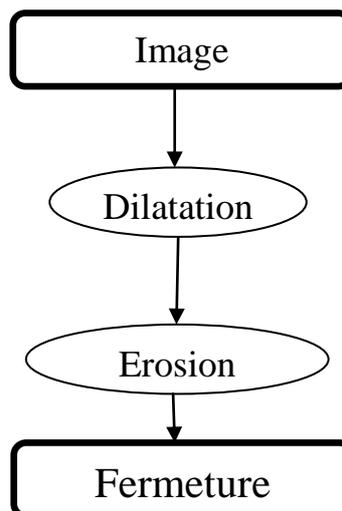


Figure II.10 : Opération de fermeture (dilatation, érosion).

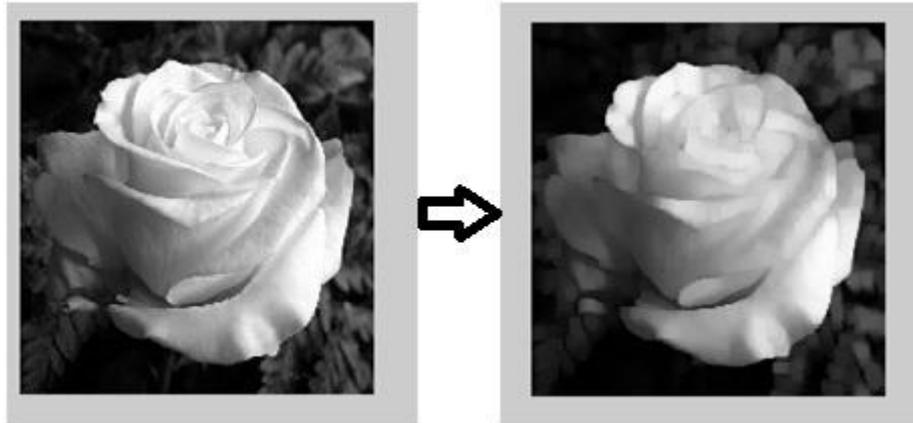


Figure II.11 : Filtre morphologique (érosion, dilatation).

II.8 Segmentation

II.8.1 Définition de la segmentation

L'opération de segmentation permet de trouver les différentes régions homogènes et ses contours, ces deux derniers sont supposés être pertinents, c'est-à-dire que les régions doivent correspondre aux parties significatives des objets du monde réel, et les contours à leurs frontières apparentes [17].

II.8.2 les méthodes de segmentation

nous pouvons résumer la méthode de segmentation dans le schéma suivant :

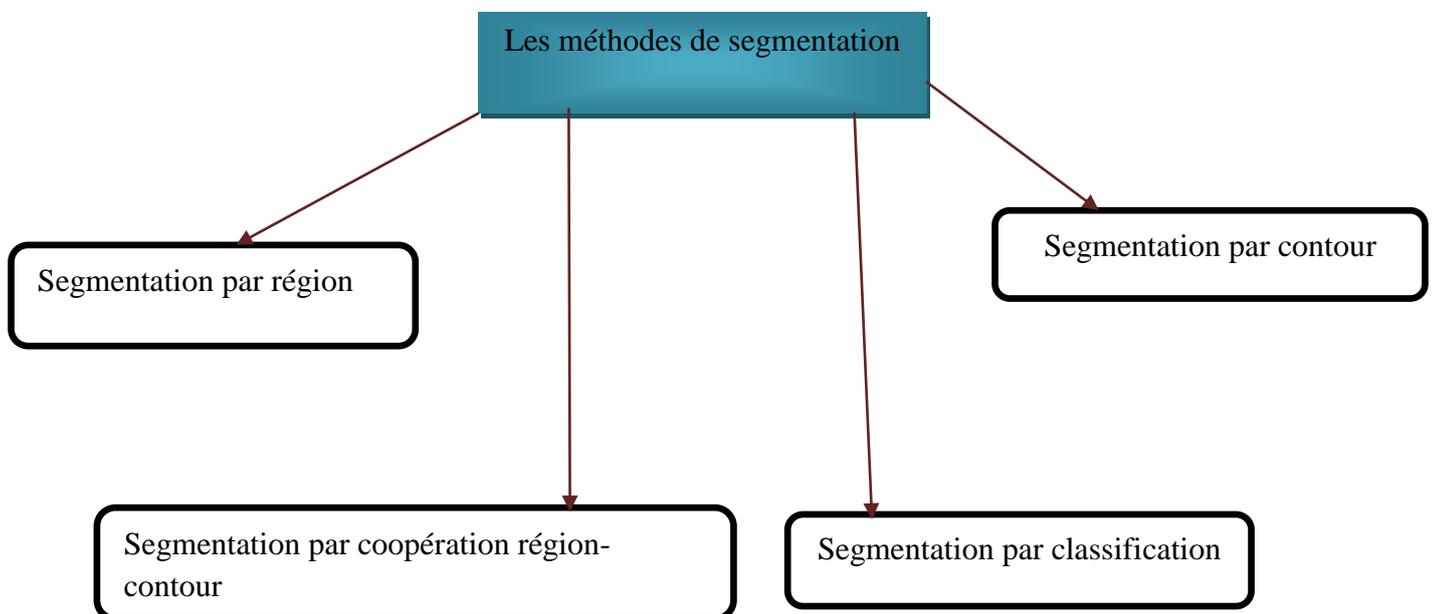


Figure II.12 : Schéma représentant les techniques de segmentation.

II.8.2.1 Segmentation par région

La segmentation d'image par l'approche région consiste à découper l'image en régions. Les pixels adjacents sont regroupés en régions distinctes selon un critère d'homogénéité ou de similarité donnée. Ce critère peut être, par exemple, le niveau de gris, couleur, texture... etc.

Un processus de groupement est répété jusqu'à ce que tous les pixels dans l'image soient inclus dans des régions. Cette approche vise, donc, à segmenter l'image en se basant sur des propriétés intrinsèques des régions [17].

II.8.2.2 Segmentation par contours

Le contour peut être pris comme le bord ou la frontière de deux régions (objets), et la détection des contours est équivalente à détecter les discontinuités à la frontière de deux régions.

Un contour est un ensemble de pixels formant une frontière entre deux ou plusieurs régions voisines, l'épaisseur d'un contour est d'un ou plusieurs pixels et il est défini par une variation "rapide" de caractéristique [18].

II.8.2.3 Segmentation par coopération région-contour

La segmentation par coopération région contour peut être exprimé comme une entre-aide entre ces deux concepts afin d'améliorer le résultat final. C'est une approche qui peut être une solution pour pallier aux inconvénients des deux approches [18].

II.8.2.4 Segmentation par classification

Dans cette section nous considérons les pixels de l'image comme étant les objets à classifier. Les techniques de la segmentation par classification fournissent une partition de l'image en regroupant des pixels ayant des niveaux de gris similaires dans une même classe de pixels. Les classes sont définies par les ensembles maximaux de pixels connexes appartenant à la même classe [18].

II.8.3 Domaine d'application de la segmentation

La technologie de traitement d'images présenté l'aspect multidisciplinaire. Nous avons trouvé ses applications dans des domaines très variés tels que :

- Médecine (Radiographies, tomographies,...) ;
- Biologie, Météorologie
- Astronomie
- Géologie
- Physique (spectroscopie, physique des plasmas,...)
- Applications militaires
- Applications industrielles (Robotique, surveillance de qualité,...)
- Photographies, Publicité, etc

II.9 Conclusion

Dans ce chapitre, nous avons décrit les différentes opérations d'une image permettant d'améliorer la qualité de celle-ci en appliquant les méthodes de prétraitement appropriées (suppression des bruits par les différents filtres,...etc.), ceci sans perdre les informations utiles de l'image. Nous avons présenté aussi la segmentation et les différentes méthodes utilisées pour cette dernière.

L'objet de chapitre suivant est de présenter différentes méthodes pour faire la reconnaissance d'objets dans les images.

CHAPITRE III :

La reconnaissance d'objets dans les images

III.1 Introduction

La reconnaissance automatique d'objets dans les images est un des problèmes les plus difficiles en vision par ordinateur. Elle est en même temps une étape primordiale pour la mise en œuvre de plusieurs applications actuelles qui nécessitent une interprétation de haut niveau d'images.

Notre objectif, dans ce chapitre est de présenter d'abord les approches et les méthodes principales de reconnaissance d'objets en général et d'évaluer leur applicabilité à notre problème.

III.2 Problème de la reconnaissance d'objets

Le problème de la reconnaissance d'objets est de décider si un objet spécifique ou un objet appartenant à une catégorie d'objets prédéfinie existe dans une image. On peut poser le problème de la reconnaissance comme un appariement entre le modèle de l'objet cible et un ensemble de descripteurs qu'on extrait à partir de l'image de test. Cette généralisation, aussi simple qu'elle semble être, est capable d'expliquer la multitude d'approches qui se différencient par le choix des descripteurs qui représentent l'objet, le type et la complexité du modèle, et les méthodes utilisées pour l'apprentissage du modèle d'objet et l'appariement.

III.3 Les principaux domaines d'application

Au delà d'être un problème général de la vision par ordinateur, la reconnaissance d'objets dans des catégories très variables est une étape indispensable pour de nombreuses applications spécifiques. On peut citer, parmi d'autres, quelques applications exemplaires pour illustrer les enjeux des travaux de recherche sur ce problème.

✓ La vidéosurveillance : la reconnaissance et la détection automatique des personnes, des visages et des objets d'intérêt sur la très grande quantité d'images vidéo utilisées pour la surveillance.

✓ La recherche dans des bases d'images numériques : les systèmes actuels effectuent en général des requêtes à partir des mots clés associés aux images. On voudrait exploiter le contenu visuel des bases d'images sans avoir besoin de les indexer textuellement.

✓ La navigation des robots mobiles et des véhicules : on a besoin de reconnaître les obstacles, les objets divers dans l'environnement ou les autres véhicules pour la navigation automatique. Pour pouvoir évaluer les différentes méthodes et pour pouvoir les comparer, il est nécessaire de créer des plateformes communes de test. Ont été définies, par des équipes de recherche qui travaillent sur ce problème, des bases d'images consacrées à la reconnaissance d'un ensemble limité de catégories d'objets. Ces bases contiennent en général un grand nombre d'images d'objets étiquetées et éventuellement des informations associées (la position, le rectangle englobant, la segmentation de l'objet). En considérant qu'elles indiquent les principales directions de recherche, nous voulons mentionner les bases d'images publiques les plus connues et les catégories d'objets répertoriées les plus communes :

- Bases d'images FERET, CMU/MIT, MIT-CBCL, Caltech, Yale : les visages ;
- Bases d'images ETHZ, Caltech, UIUC : les voitures, les motocyclettes ;
- Bases d'images Pascal, CMU : les bicyclettes ;

- Bases d'images Graz 02, Pascal : les personnes, les vaches ;
- Base d'images Weizmann : les chevaux ;

III.4 La vision d'être vivants

L'œil humain est capable de reconnaître rapidement une multitude d'objets différents avec des variations de point de vue, d'illumination et de forme, même dans les contextes les plus inhabituels. On reconnaît les catégories d'objets aussi bien que les instances d'objets uniques, et cela avec une capacité de généralisation sur plusieurs échelles et sous plusieurs sortes de transformations. On peut ajouter à cela l'apprentissage rapide d'objets qu'on voit pour la première fois.

On est loin de connaître le modèle complet de la perception humaine qui réunit des mécanismes cognitifs complexes et de l'imiter pour développer un système automatique capable d'approcher la performance humaine. Pourtant, les expériences psycho-cognitives sur les étapes primales (la reconnaissance rapide sans utiliser les processus cognitifs de haut niveau) de la vision humaine et sur la vision des pigeons ont révélé des indices qui peuvent être exploités par les systèmes automatiques. Il nous semble pertinent de citer deux travaux expérimentaux importants.

III.4.1 Théorie des caractéristiques particulières

Cerella propose que la vision chez les pigeons fonctionne comme des perceptrons. Chaque perceptron correspond à une caractéristique simple locale de l'objet et la reconnaissance est liée aux réponses des perceptrons individuels. Pour justifier sa théorie, il présente une expérience où les pigeons reconnaissent de la même façon un objet quand ses composants sont mélangés. Cela montre que ce qui est prioritaire pour la reconnaissance d'objets est la détection des caractéristiques locales mais non pas l'organisation spatiale des primitives [19].

III.4.2 La théorie de la reconnaissance par composants

Venkateswar a développé la théorie de la reconnaissance par composants pour expliquer la reconnaissance d'objets par les humains. Selon sa théorie, les objets sont constitués de composants primitifs et on les reconnaît en détectant ces composants et leurs relations.

Les composants élémentaires sont appelés "*les géons*" (geometrical ions) [20].

Ce sont des formes géométriques tridimensionnelles, obtenues par la déformation d'un cylindre (Figure III.1), avec les propriétés suivantes :

- Invariantes au point de vue : ils sont identifiables de plusieurs points de vue différents.
- Discriminantes : il est difficile de confondre un géon avec un autre.
- Résistantes au bruit : on peut les détecter même en cas d'imperfection de la scène.

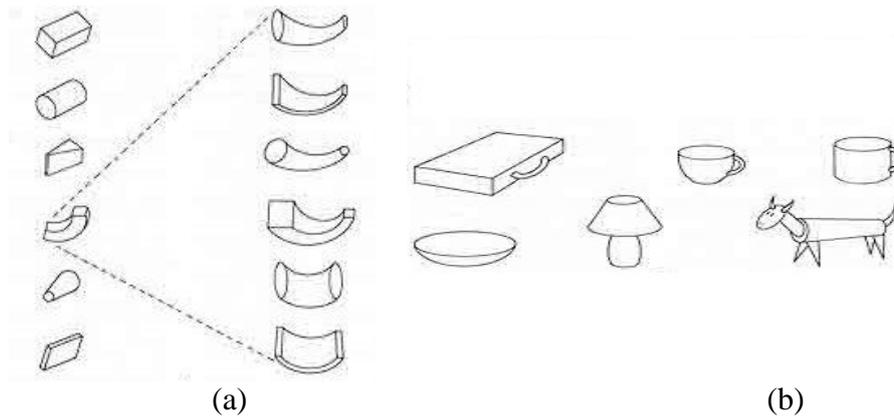


Figure III.1 : (a). Exemples de géons, (b). Quelques objets composés de géons.

III.5 Les défis majeurs de la reconnaissance d'objets

La difficulté majeure du problème de la reconnaissance d'objets repose sur les variations d'un objet ou des objets d'une même catégorie sur des images différentes. On peut grouper les sources de variations dans sept catégories principales :

- **Les variations de point de vue** : les objets réels sont des entités physiques tridimensionnelles. Leur apparence change de façon significative en fonction du point de vue.
- **Les variations d'illumination** : en fonction de changements de l'environnement et des conditions d'acquisition d'une image, un objet peut subir des variations d'illumination importantes. Les objets dans les images des scènes naturelles sont très sensibles aux variations d'illumination qui modifient les couleurs et les valeurs radioélectriques des objets considérablement.
- **Les occlusions** : dans une image les objets peuvent être partiellement masqués par d'autres objets ou par leurs propres composants.
- **Les variations d'échelle** : les objets d'une même catégorie peuvent avoir des tailles physiques différentes ou les images d'un même objet peuvent avoir des tailles différentes en fonction de la résolution de l'image.
- **Les déformations** : les pluparts des objets réels n'ont pas une forme rigide. Particulièrement les objets articulés subissent généralement des déformations linéaires et non-linéaires de leurs composants.
- **Complexité du fond** : en général, dans une image les objets ne sont pas isolés de leur environnement. Sur un fond complexe il est difficile de distinguer les régions appartenant à l'objet et au fond.
- **Variations intra-classe** : c'est le plus important des problèmes envisagés pour la reconnaissance générique d'objets. Les objets d'une même classe peuvent avoir des apparences très variables, même quand on se limite aux catégories à un niveau élémentaire. Dans le cas extrême, les catégories fonctionnelles comme la classe "chaise" sont définies par un haut niveau d'abstraction et les indices visuels seuls ne sont pas suffisants pour les reconnaître.

L'objectif des méthodes de reconnaissance d'objets est de trouver des descripteurs et des représentations qui soient invariants aux variations de l'objet présentées ci-dessus. Cet objectif étant très ambitieux, pendant la phase initiale de la définition du problème de reconnaissance et de la construction de la base d'images on applique souvent des contraintes qui limitent les éventuelles variations des objets.

Les images satellites offrent un point de vue constant. Elles n'ont pas des variations d'illumination importantes. Les objets cartographiques ne sont pas en général sujets à des occlusions.

Dans notre base toutes les images ont la même résolution spatiale. Pourtant les objets ont des tailles physiques assez différentes. Ce sont des objets articulés qui ont subi des déformations importantes et qui ont une grande variation intra-classe. Finalement il est très difficile de séparer les objets du fond à cause de la ressemblance des composants d'objets cibles à des autres objets environnants. Ces considérations ont eu un impact crucial pour la sélection des méthodes que nous avons proposées.

III.6 La reconnaissance spécifique versus générique

On peut faire la distinction entre deux problèmes de reconnaissance différents : la reconnaissance Spécifique on cherche à détecter un objet précis qui a éventuellement subi des déformations perspectives ou des transformations d'illumination. Les exemples sont la reconnaissance d'un jouet d'enfant, la reconnaissance d'un produit spécifique d'une marque, la reconnaissance des panneaux de signalisation ou l'identification d'une personne. Dans la reconnaissance générique, par contre, l'objectif est de détecter les objets appartenant à une certaine catégorie. Ce problème est considérablement plus difficile que le précédent. La catégorie d'un objet n'est pas un attribut intrinsèque de l'objet, elle est souvent apprise par l'observation des similarités et des différences d'objets distincts. L'homme est capable d'attribuer une catégorie à un objet vu pour la première fois et aussi de créer des nouvelles catégories pour de nouveaux objets. Par conséquent, les objets dans une même catégorie peuvent avoir des apparences très variables et la difficulté majeure du problème de la reconnaissance générale est de trouver une description de la catégorie qui permet de représenter ces variances. [21]

III.7 Le schéma général des méthodes de reconnaissance d'objets

1. **Extraction** : l'extraction d'un ensemble de descripteurs à partir de l'image, qui reflètent les informations pertinentes pour la reconnaissance de l'objet ;
2. **Représentation** : la construction d'un modèle d'objet pour représenter une catégorie d'objet;
3. **Apprentissage** : l'apprentissage des paramètres du modèle à partir des images d'apprentissage étiquetées ;
4. **Reconnaissance** : l'appariement avec le modèle ou la classification en utilisant le modèle pour reconnaître l'objet cible dans une image de test.

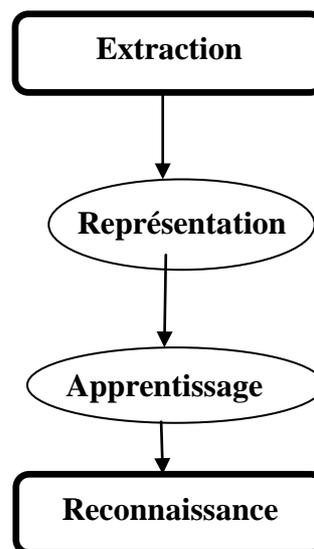


Figure III.2 : Le schéma général des méthodes de reconnaissance d'objets

- ❖ Les méthodes de reconnaissance d'objets se différencient principalement au niveau de :
 - La définition du problème de reconnaissance : les catégories d'objets recherchées, le nombre de catégories.
 - Le nombre d'images utilisées pour l'apprentissage : une extrémité étant l'apprentissage à partir d'une seule image, on peut aussi utiliser un très grand nombre d'images pour l'apprentissage.
 - Le niveau de supervision : en cas d'un apprentissage non-supervisé on ne dispose pas d'images étiquetées, et on essaie de grouper les images dans des catégories d'après leurs similarités (clustering). En cas d'apprentissage supervisé, on a un ensemble d'images d'apprentissage étiquetées. Il s'agit d'un niveau de supervision plus élevé quand on connaît aussi la segmentation manuelle des parties de l'objet sur l'image.
 - La modélisation : ici, nous utilisons le terme modèle dans le sens général de "la représentation d'un objet". Un très grand nombre de modèles d'objets différents ont été proposés pour représenter les objets dans les images.
 - Les méthodes et les algorithmes choisis pour l'extraction des descripteurs, la construction du modèle et la reconnaissance.

La reconnaissance d'objets nécessitant une chaîne de traitement qui se constitue de plusieurs étapes, il est difficile de présenter la multitude d'approches utilisées sans analyser chacun des principaux composants individuellement. Même si les traitements de haut niveau ont une grande importance pour la reconnaissance, elles se sont construites sur les résultats d'une étape de traitement de bas niveau qui est également importante et toujours d'actualité en vision par ordinateur. Nous introduisons ci-dessous les approches existantes pour ces étapes majeures, avant de présenter les méthodes de reconnaissance d'objets les plus significatives.

III.8 Les traitements d'extraction des descripteurs

L'utilisation directe de l'image pour la reconnaissance n'est pas envisageable parce qu'elle contient une grande quantité d'information impertinente ou redondante. C'est pourquoi, la première étape de presque toutes les méthodes de reconnaissance est d'extraire un ensemble de descripteurs qui représentent les propriétés discriminantes de l'objet cible. Les descripteurs sont un ensemble de nombres scalaires générés pour décrire un objet (ou une forme). Un objet ne peut pas être reconstruit en général à partir des descripteurs, mais les descripteurs pour des objets distincts doivent avoir des valeurs suffisamment différentes pour les discriminer, et les descripteurs pour les objets similaires doivent avoir des valeurs proches. On peut extraire un nombre infini de descripteurs à partir d'une image, pourtant l'objectif est d'en sélectionner un nombre moins important que le nombre de pixels dans l'image. Pour une reconnaissance robuste il est souhaitable que les descripteurs soient invariants aux changements d'échelle et d'illumination et aux transformations affines.

On peut distinguer deux types de descripteurs : **les descripteurs globaux** sont obtenus à partir de l'image entière. Par contre, **les descripteurs locaux** caractérisent un ensemble de régions locales significatives de l'image qui sont riches en information. Les travaux récents en reconnaissance d'objets se concentrent le plus souvent sur l'extraction des descripteurs locaux qui sont plus efficaces pour la reconnaissance générique. On utilise principalement les contours, les régions ou finalement les propriétés de texture et de couleur de l'image pour extraire les descripteurs. Ci-dessous, nous présentons les principaux types de descripteurs et les méthodes utilisées pour les extraire.

III.8.1 Les points d'intérêt

Les points d'intérêt sont des régions de l'image riches en termes de contenu de l'information locale et stables sous des transformations affines et des variations d'illumination. Ils sont des indicateurs des régions susceptibles de contenir un objet, et en même temps des parties importantes de l'objet. Les points d'intérêts sont plus stables que les régions ou les contours de l'image et leur extraction est plus simple.

III.8.1.1 Détection des coins

Le détecteur des coins de *Harris* est l'un des plus populaires des détecteurs de points d'intérêt. Un coin, l'intersection de deux contours, a un contenu informatif important. Sur l'image les coins correspondent aux pixels qui ont des changements d'intensité élevés dans toutes les directions sur un voisinage local (par contre sur les régions plates le changement d'intensité est nul dans chaque direction, et sur les contours il est élevé dans une seule direction). L'idée principale est donc de trouver les points à partir desquels un petit

déplacement donne un changement d'intensité élevée. Pour cela, on calcule M , la matrice de covariance du gradient dans une fenêtre de $2n + 1$ autour d'un pixel p . Les deux valeurs propres de M correspondent au changement minimal et au changement maximal d'intensité au voisinage du p . Deux valeurs propres qui sont supérieures à un seuil (précisé empiriquement) indiquent l'existence d'un coin. La figure III.3 montre les coins détectés sur une image satellitaire.

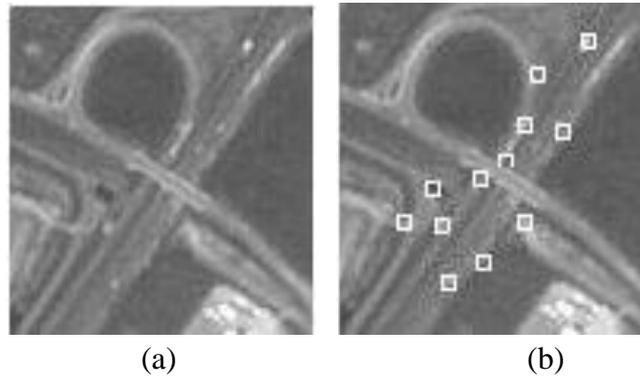


Figure III.3: Les coins détectés sur une image satellitaire par le détecteur [22]

Les coins de *Harris* sont très sensibles au changement d'échelles et ils ne contiennent pas d'information sur la région locale autour du point d'intérêt. Par conséquent il n'est pas possible d'apparier correctement les points d'intérêt correspondants sur deux images sauf si les objets sur les images ont des formes précises et identiques [22].

III.8.1.2 Les descripteurs SIFT

Schmid ont proposé de calculer un descripteur local pour chaque point d'intérêt. Le descripteur local est calculé à l'aide des invariants différentiels de l'intensité des pixels sur un voisinage local autour du point d'intérêt. Pour obtenir l'invariance à l'échelle on utilise un ensemble de dérivées Gaussiennes avec des valeurs de différentes. La figure III.4 illustre l'approche utilisée. L'association des descripteurs locaux aux points d'intérêt rend possible un nombre d'appariement correct élevé entre deux images qui contiennent le même objet. Cela assure la robustesse de la reconnaissance d'objets [23].

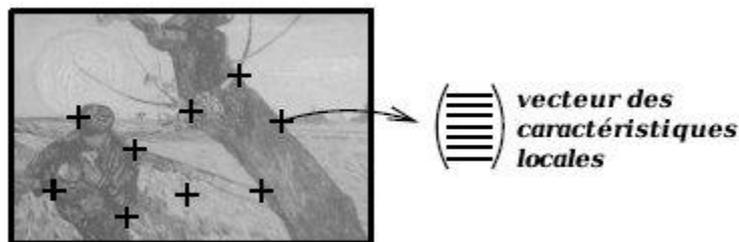


Figure III.4 : Les points d'intérêt et les descripteurs locaux [23]

La méthode SIFT (Scale Invariant Feature Transform) est basée sur la même idée. C'est une méthode actuellement très populaire qui détecte des points d'intérêt et leurs descripteurs locaux invariants à l'échelle et aux rotations, et invariants partiellement aux changements d'illumination. Les descripteurs obtenus sont très distinctifs. On peut apparier un descripteur

correctement avec une probabilité très élevée parmi une grande base de descripteurs obtenus d'autres images.

L'algorithme procède en quatre étapes principales : la détection des extrema (qui correspondent aux points d'intérêt) a un espace d'échelle créé par des filtres DoG sur une pyramide multi échelle ; la localisation et le filtrage des points d'intérêt ; le calcul des histogrammes des directions des gradients locaux autour du point d'intérêt dans un voisinage qui varie en fonction de l'échelle du point, et le calcul d'une signature à partir des histogrammes en divisant la zone autour du point en quatre sous-zones. La signature est un code de taille 32 qui indique la répartition des 8 orientations principales pour chaque zone (Figure III.5) [24].

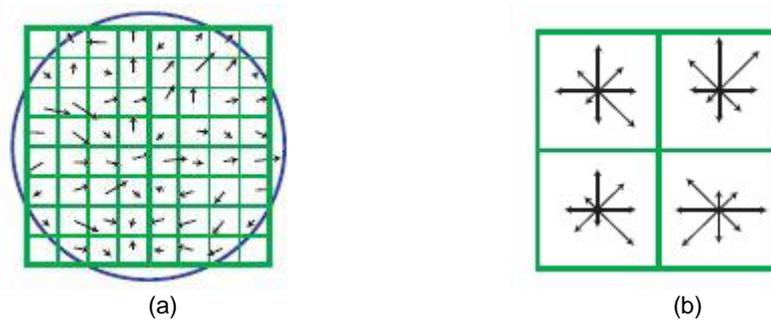


Figure III.5 : SIFT, calcul de la signature d'un point d'intérêt. (a). Les gradients de l'image autour du point d'intérêt, (b). Les histogrammes correspondants

L'objet recherché est représenté finalement par un ensemble de points d'intérêt et les signatures associées, et la reconnaissance est effectuée par un appariement des points d'intérêt entre les images. Grâce aux propriétés d'invariance et à la haute capacité descriptive des points d'intérêt, les résultats sont excellents pour la reconnaissance d'un objet spécifique dans une autre image. L'appariement de seulement trois points suffit pour une détection fiable. Pourtant, les méthodes de reconnaissance par la correspondance des points d'intérêts ne sont pas adaptées pour la reconnaissance générique surtout quand la variation intra-classe est élevée. Les descripteurs locaux sont trop spécifiques pour apparier deux objets dans la même catégorie mais avec des variations. Pour notre problème, les objets appartenant à une classe n'ont pas de caractéristiques locales similaires, mais ont des caractéristiques communes structurelles à un plus haut niveau d'abstraction. Pour cette raison, nous n'avons pas décidé d'utiliser ces méthodes.

III.8.1.3 Les blobs

Les blobs sont des régions de l'image qui sont plus claires ou plus sombres que leurs environs. On peut les considérer comme des points d'intérêt qui signalent les objets ou les composants des objets. Le détecteur le plus fréquemment utilisé est le filtre LoG qui donne des valeurs positives élevées pour les blobs sombres et des valeurs négatives élevées pour les blobs claires. Pourtant la réponse de l'opérateur dépend de l'échelle du LoG. *Lindeberg* a proposé d'utiliser un opérateur Laplacien normalisé sur un espace d'échelle et de trouver les maxima locaux dans l'espace et dans l'échelle simultanément. De cette façon on détecte les blobs à l'échelle correspondante [25].

III.8.1.4 Régions extrêmes maximales stables (MSER)

Les MSER est une technique pour détecter les blobs dans une image. On obtient une série d'images noires et blanches en appliquant un seuillage pour chaque niveau de gris dans l'image. Les MSER sont les blobs qui restent stables sur plusieurs images consécutives.

III.8.2 Les descripteurs contours

Les contours d'un objet donnent suffisamment d'information sur sa forme. D'après la définition du dictionnaire le mot "forme" signifie :

Forme : Ensemble des contours d'un objet, d'un être, résultant de la structure de ses parties.

L'œil humain est capable de reconnaître un objet à partir des fragments de ses contours (Figure III.6).

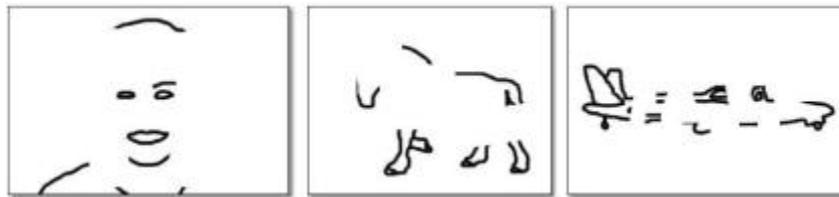


Figure III.6 : Les fragments de contours de quelques objets courants.

Les descripteurs obtenus à partir des contours ont l'avantage d'être invariants aux changements d'intensité, de texture et de couleur de l'image. Les descripteurs contours sont des candidats idéaux pour représenter les propriétés structurelles des objets.

III.8.2.1 Détection des contours

Les contours correspondent aux changements brusques de l'intensité de la lumière sur l'image. Sur une image numérique ces changements peuvent être détectés à l'aide des dérivées du premier et du second ordre de la fonction d'intensité. Un point de contour est défini comme le maximum du module du gradient dans la direction du gradient. Avant la détection des contours on procède à des prétraitements pour supprimer le bruit provenant des imperfections des capteurs. Le filtre le plus fréquemment utilisé pour supprimer le bruit est le filtre gaussien $G_{\sigma}(x, y)$ qui pondère la valeur de chaque pixel selon une distribution normale autour de ce pixel. Cela résulte en un lissage de l'intensité de l'image. La dimension du filtre dépend du σ et plus on a un σ élevé, plus le lissage est important. En convoluant l'image par une série de filtres gaussiens avec des valeurs de σ croissantes, on peut obtenir une représentation multi échelle de l'image qu'on appelle la pyramide gaussienne.

Une méthode de filtrage adaptatif est la diffusion anisotropique où on contrôle le niveau de la diffusion, donc du lissage, par un paramètre d'échelle qui la réduit quand l'intensité du gradient augmente [26].

La détection est souvent suivie d'une étape de post-traitement qui sert à réduire le nombre de points de contours et à les relier. Très souvent, on élimine les points de contour qui ont un gradient faible par un seuillage. Des algorithmes de suivi de contours peuvent ensuite être appliqués pour connecter les points de contours adjacents.

Le détecteur de Canny : Le détecteur de Canny est considéré par plusieurs comme le détecteur optimal. Canny a précisé trois critères pour un bon détecteur et a proposé une

méthode qui satisfait ces trois critères. Ce sont : la bonne détection (faible taux d'erreur), la bonne localisation (minimisation des distances entre les contours réels et contours détectés) et la clarté de la réponse (une seule réponse par contour) [27].

Les modèles déformables : Une autre approche de détection, très souvent utilisée pour les applications médicales, est la méthode des modèles déformables. L'idée principale est de faire évoluer itérativement une courbe initiale afin qu'elle s'approche à chaque itération du contour réel jusqu'à ce qu'elle reste constante. L'évolution de la courbe est faite par la minimisation d'une fonction d'énergie qui est la somme d'énergies internes et externes. L'énergie interne assure la régularisation de la forme du contour et l'énergie externe l'attache au contour réel (aux valeurs élevées du gradient). Les deux grandes familles des modèles déformables sont les contours actifs (snakes) où on utilise un modèle paramétrique pour représenter la courbe, et les ensembles de niveau (level sets) où on utilise un modèle géométrique. Ces méthodes sont très sensibles à l'initialisation de la courbe et sont mieux adaptées quand on dispose de connaissances a priori sur la forme de l'objet recherché [28].

III.8.2.2 Extraction des descripteurs contours

Les contours d'un objet ne sont pas utilisables directement pour représenter l'objet mais on procède en général à une étape d'extraction des descripteurs.

III.8.2.2.1 Les fragments de contours

Shotton propose d'utiliser des fragments de contours comme descripteurs. A partir d'un petit nombre d'images d'objets segmentées on construit un dictionnaire des fragments. Le dictionnaire est constitué des fragments des contours extérieurs de l'objet et des positions des fragments par rapport au centre de l'objet. Pour construire le dictionnaire, on génère aléatoirement un grand nombre de régions qui contiennent une partie de l'objet et du fond. On perturbe légèrement ces régions par des transformations aléatoires pour obtenir un dictionnaire plus représentatif, et on détecte les contours sur chaque région [29].

Ferrari représenté un objet par un réseau de segments de droites. Après l'extraction des contours, on procède à une approximation polygonale et un chaînage des segments par une recherche locale à leurs extrémités. On définit un ensemble de règles pour connecter les contours. On représenté finalement les contours de l'objet par un graphe de segments de contour. Suivant la même idée principale, *Ferrari* apporte une amélioration majeure en proposant un codage (par un descripteur numérique) des propriétés géométriques de chaque groupe de k segments adjacents (kAS). Le codage est invariant à l'échelle, aux translations et aux rotations. Le nombre de kAS obtenus étant très élevé pour l'appariement, on construit un dictionnaire de codes visuels () en groupant les kAS similaires par l'algorithme de partition des cliques. Au delà de $k = 3$ les kAS représentent des structures très spécifiques et leur nombre devient très important. Pratiquement on construit le dictionnaire de codes pour $k = 2$ et $k = 3$ (Figure III.7) [30].

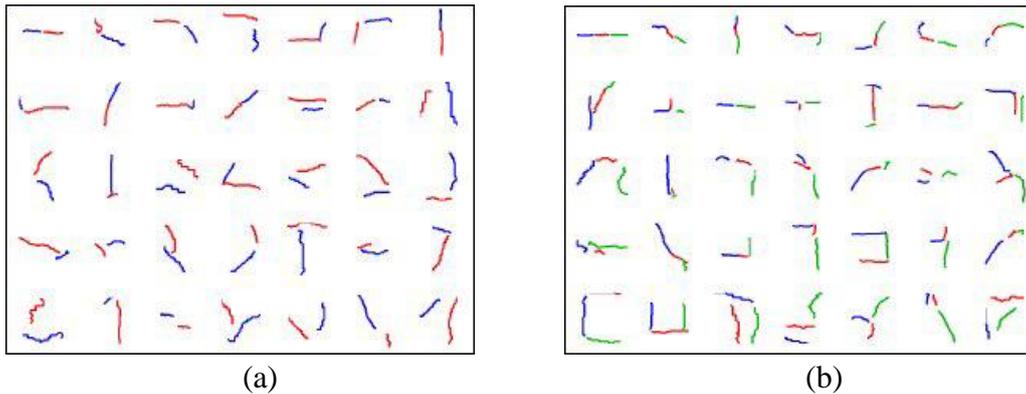


Figure III.7 Les 35 kAS les plus fréquents du dictionnaire de codes

Les kAS ont les bonnes qualités de couvrir des chaînes de contours très courtes, d'avoir une complexité intermédiaire, d'être basés sur la connexité des segments (les segments connexes ont plus de chance de se trouver sur les frontières d'un objet) et d'avoir des descripteurs invariants localement. Il est intéressant de voir que les dictionnaires de codes construits à partir d'images différentes sont assez similaires. Cela sert à montrer la généralité des kAS.

III.8.2.2 Les primitives géométriques

Une alternative pour obtenir des descripteurs à partir des contours est de rechercher des formes qui ont des propriétés géométriques bien définies. Les objets construits par l'homme sont en général constitués de composants qui ont des formes géométriques simples, contrairement aux objets naturels qui ont des formes plus irrégulières. Les formes géométriques les plus générales sont les segments de droites, les rectangles et les cercles.

La transformée de *Hough* est une méthode pour détecter les formes simples à partir d'un ensemble de points. L'application la plus classique est la détection des segments de droites. Le principe est de représenter les segments entre chaque paire de points dans un espace de coordonnées défini par l'angle et la distance du segment à l'origine et d'accumuler les nouvelles coordonnées dans une matrice. Les maxima locaux de cette matrice indiquent les segments. On peut appliquer la transformée de *Hough* également pour détecter les cercles. Cette fois-ci on fait la transformation vers un espace à trois coordonnées, position x , position y et le rayon. La recherche des cercles dans cet espace est difficile surtout quand on ne connaît pas les rayons des cercles recherchés [31].

La transformée de *Hough* traite les contours comme un ensemble de points indépendants. A la fin de l'étape de l'extraction des contours on dispose souvent des informations sur les connections entre les points de contours et la direction du gradient sur ces points. On peut exploiter ces informations pour une meilleure extraction des primitives.

Venkateswar ont détecté les segments de droites dans les images aériennes en parcourant l'image pour affecter une étiquette de segment à chaque point de contour d'après la direction du gradient à ce point. On regroupe ensuite les points qui ont les mêmes étiquettes et on détermine les extrémités [32]. Dans une méthode similaire, *Desolneux* utilisent le principe d'évènement significatif : dans une image, les directions des gradients de deux points voisins sont des variables aléatoires. Les déviations de ce caractère aléatoire sont considérées comme

des événements significatifs qui correspondent aux structures. Ils ont calculé la direction du gradient très localement, dans un voisinage de 2×2 et recherché les événements significatifs. Ils ont utilisé leur méthode pour détecter les structures linéaires dans les images (Figure III.8). Une autre approche est l'utilisation des algorithmes d'approximation polygonale qui transforment une chaîne de contours en une suite de segments de droites [33]. *Garnesson* présentent une vue d'ensemble du problème et les méthodes utilisées. Deux algorithmes polygonaux classiques par découpage sont l'algorithme de la corde *Douglas et Peucker*. On commence par un segment grossier qui représente la chaîne de contour, on trouve le point de découpage (qui satisfait les critères spécifiques à l'algorithme) et on divise le segment en deux par ce point. On répète cette étape récursivement jusqu'à ce qu'on obtienne des segments avec une erreur d'approximation acceptable [34].

Iqbal ont proposé d'extraire d'autres primitives par les principes de regroupement perceptuel des segments de droites, la proximité, la similarité, la continuation, la fermeture et la symétrie. Les primitives obtenus sont les coterminations, les jonctions L et T (Les jonctions L sont les paires de segments qui s'articulent à leurs extrémités et les jonctions T sont les paires de segments où l'articulation est entre l'extrémité d'un segment et le milieu de l'autre), les segments parallèles, les groupes parallèles et les polygones [35].

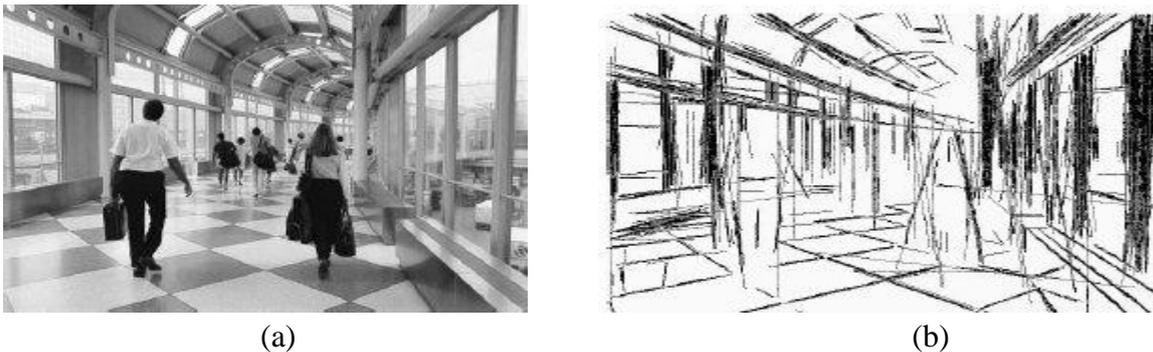


Figure III.8 : Les segments linéaires par la détection des événements significatifs

III.8.3 Les descripteurs

L'objectif de ces approches est de détecter dans l'image les régions qui appartiennent à l'objet et qui ont une forme particulière. On peut ensuite représenter les régions détectées par un ensemble de descripteurs de formes.

III.8.3.1 Les descripteurs de formes

Une fois l'image segmentée, on peut représenter les régions obtenues par des descripteurs de formes. On représente une forme comme une région ou par ses contours pour extraire les descripteurs.

III.8.3.1.1 Les descripteurs obtenus par une approche région

Pour calculer ces descripteurs la forme est considérée comme l'ensemble des pixels qui constituent sa région interne.

Les descripteurs géométriques simples :

- L'aire, le périmètre.
- La compacité ou la circularité : le rapport du carré du périmètre à l'aire de la forme

La forme la plus compacte est un cercle avec une compacité égale à 1.

- La rectangularité : c'est le rapport de l'aire de la forme à l'aire du plus petit rectangle englobant.
- L'excentricité : la mesure de la circularité d'une ellipse. Pour une ellipse c'est le rapport de la distance entre les foyers à la longueur de l'axe majeur. On calcule l'excentricité de l'ellipse qui a les mêmes moments d'ordre deux que la forme.
- L'elongation : le rapport de la longueur à la largeur du plus petit rectangle englobant.
- Les axes principaux : l'axe majeur et l'axe mineur de l'ellipse qui a les mêmes moments d'ordre deux que la forme.
- L'orientation : la direction de l'axe majeur.

Les descripteurs topologiques : Deux formes sont topologiquement identiques si on peut obtenir une en déformant l'autre par des transformations continues. Les descripteurs topologiques de base sont le nombre de composants connexes et les trous. Le plus connu de ces descripteurs est le nombre d'Euler qui est égal au nombre de composants connexes moins le nombre de trous.

L'enveloppe convexe : Une région R est convexe si pour chaque paire de points p_1, p_2 de R le segment $[p_1p_2]$ est dans R . L'enveloppe convexe est la plus petite région convexe qui contient la forme.

Les profils : On peut obtenir une signature de l'objet en le projetant sur l'axe horizontal ou vertical et en comptant le nombre de pixels sur chaque colonne ou ligne. On utilise ces descripteurs le plus souvent pour la reconnaissance des caractères [36].

Les moments : La fonction d'intensité de la forme peut être représentée comme une densité de probabilité d'un variable aléatoire bidimensionnelle. Les moments géométriques décrivent les caractéristiques statistiques de cette variable aléatoire [37].

Le squelette : Le squelette de l'objet est un ensemble de segments connexes d'épaisseur un pixel obtenu par l'amincissement d'une forme. On l'appelle également l'axe médian, l'ensemble des points qui ont la même distance minimale aux frontières de la forme pour au moins deux points de frontière distincts. On peut obtenir le squelette d'une forme par les opérateurs de la morphologie mathématique ou par une transformée de distance qui associe à chaque pixel la valeur de la distance minimale à la frontière.

III.8.3.1.2 Les descripteurs obtenus par une approche contours

Ce sont des descripteurs obtenus à partir des contours externes de l'objet.

Les descripteurs géométriques simples :

- Le périmètre.
- La courbure : le contour est représenté par l'angle de la tangente à chaque point de contour. On peut ensuite calculer la dérivée des courbures sur chaque point de contour ou on peut détecter les courbures importantes du contour comme descripteur.
- L'énergie de flexion (*bending*) : la somme des carrés des courbures à chaque point de contour.
- La distribution des cordes : une corde est un segment de droite qui joint deux points de contours. On peut mesurer la distribution des cordes par les histogrammes des longueurs et des angles de toutes les cordes du contour.

Le codage de chaîne : A partir d'un point de contour initial, on traverse le contour en codant les points consécutifs par la direction du déplacement. Les chiffres de 0 à 7 sont utilisés pour coder les 8 directions principales. Le codage est plus économique que représenter le contour par les coordonnées des points de contours, et il est invariant à la translation. Pour assurer l'invariance à la rotation on utilise le codage de chaîne différentiel en calculant les différences des directions des déplacements consécutifs en mode 8. La sélection du point de départ qui donne le code avec la plus petite valeur numérique rend le code invariant à la sélection du point de départ.

Approximation polygonale : Les contours de la forme peuvent être représentés par des segments de droites obtenus par une approximation polygonale

III.8.3.2 Les descripteurs de texture et de couleur

Les descripteurs de textures sont utilisés dans plusieurs domaines du traitement d'image, comme l'imagerie médicale, le traitement des documents et la télédétection. Contrairement à la couleur, la texture n'est pas la propriété d'un pixel mais d'une région. Définir exactement ce qu'est la texture n'est pas facile. Par conséquent, il existe un nombre très important de travaux pour définir, segmenter et classifier les textures, groupés en 4 approches principales. Dans l'approche structurelle, on distingue deux éléments principaux, les primitives à utiliser et les relations spatiales qui lient les primitives entre elles. Dans l'approche statistique, la texture est vue comme la réalisation d'un processus stochastique. Le but est alors d'en extraire des attributs statistiques.

III.9 Les traitements de haut niveau : reconnaissance d'objets

Au début des travaux de recherche en reconnaissance d'objets dans les années 1970, les représentations géométriques ont joué un rôle important. On a souvent ciblé la reconnaissance d'objets rigides extraite à partir des contours de l'image, a fourni une base théorique solide pour gérer les variations d'illumination et de perspective de ces objets

III.9.1 L'apprentissage structurel

L'interprétation d'images (image understanding) est un domaine de recherche devenu très populaire dans les années 90, qui unit deux disciplines proches, le traitement d'images et l'intelligence artificielle. Elle vise principalement à associer les régions d'une image numérique à des objets sémantiques pour obtenir une interprétation de haut niveau du contenu d'une image en essayant d'intégrer des connaissances externes (représentées par exemple par des ontologies) dans le traitement. La reconnaissance de formes structurelle est souvent vue comme un moyen de combler le fossé sémantique entre les descripteurs de bas niveau et les concepts de haut niveau. Elle est caractérisée par la représentation des formes en termes de structures de données symboliques pour révéler une description de la structure d'un objet. Cette description est en général basée sur les attributs des composants de l'objet et leurs relations dans l'espace représentés par un modèle structurel d'objet, et la méthode principale de la reconnaissance de formes structurelle est l'appariement des structures [38].

On effectue la modélisation souvent pour des problèmes jouets avec des objets très simples ou synthétiques (un exemple est donné dans la figure III.9). Les modèles obtenus sont simples et souvent trop rigides pour permettre la reconnaissance sur les images réels.

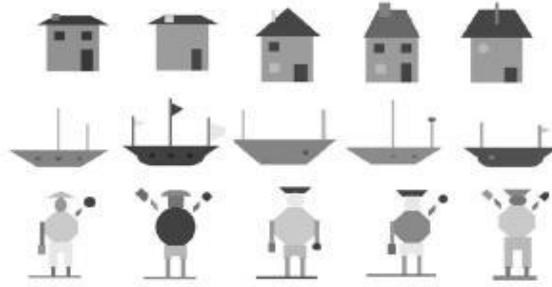


Figure III.9 : Les images d'apprentissage artificielles créées par une grammaire

III.9.2 L'approche fondée sur l'apparence

On peut utiliser des descripteurs globaux ou locaux pour révéler l'apparence d'un objet sur une image. La détection des visages en utilisant les *eigenfaces* est l'approche globale la plus remarquable [39]. L'idée principale est de projeter les images de visages sur l'espace des caractéristiques qui encode mieux les variations entre les images. Cet espace est défini par les *eigenfaces*, les vecteurs propres de l'ensemble des images de visages. On calcule les *eigenfaces* par l'analyse en composantes principales (ACP), une méthode de réduction de dimension bien connue en analyse de données.

Les approches globales, bien qu'elles soient rapides, ne sont pas adaptées pour la reconnaissance des objets non-rigides qui ont subi des déformations, et des objets sur des fonds complexes. Les méthodes récentes sont fondées sur une approche locale, le plus souvent sur les régions locales significatives de l'image dont l'apparence reste stable sur plusieurs images du même objet. Un objet est souvent conçu comme un ensemble de ses composants qui peuvent être décrits par des caractéristiques (Figure III.11).



Figure III.10 – La conception d'objet dans les approches basées sur l'apparence.

III.9.3 Les modèles par l'apparence seule

Un modèle efficace pour la reconnaissance générique d'objets. C'est un modèle initialement utilisé pour la classification des textes où on représente un texte uniquement par les mesures statistiques du vocabulaire utilisé. On construit un vecteur de caractéristiques en comptant les mots dans le texte sans prendre en compte ni la grammaire ni l'ordre des mots. Pour deux textes de même catégorie la probabilité de trouver les mêmes mots sont élevés et par conséquent les vecteurs de caractéristiques sont similaires.

III.10 La reconnaissance d'objets sur les images satellites et aériennes

En ce qui concerne les images satellitaires, la plupart des travaux aboutis concernent l'extraction de structures étendues (linéaires comme les réseaux routiers ou surfaciques comme les zones urbaines). Cette étape d'analyse des images satellitaires correspondait à un niveau de résolution peu élevé où les structures devenaient des textures [40].

Presque toutes les méthodes de détection de routes suivent la même approche : la détection des fragments des routes évidents et le suivi de ces fragments pour construire le réseau routier.

Les étapes du système sont la détection de points d'amorces (les fragments de routes évidents) à partir des contours de l'image, le suivi des points d'amorces et la reconstruction des fragments de routes qui manquent d'après un certain nombre de règles de connexions. Pourtant, il a limité son travail à la détection des carrefours simples à branches droites. On utilise une approche heuristique pour la détection des carrefours : on calcule des profils radiométriques concentriques autour d'un point candidat qui sont discriminants pour les carrefours.

L'auteur de [41] a travaillé sur le problème spécifique de la détection des ronds-points sur les images satellites. On a défini deux descripteurs locaux, une circulaire qui analyse les pixels sur des cercles autour du point et un autre radial sur des rayons (Figure III.11) [41].

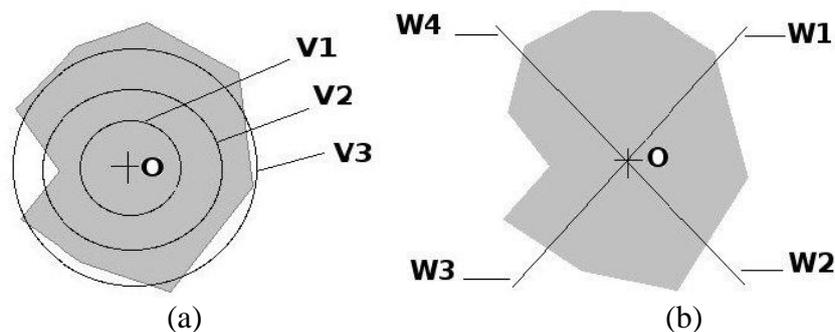


Figure III.11 : Les descripteurs radiaux et angulaires

Les auteurs de [42] ont traité le sujet de reconnaissance d'objets construits par l'homme. Les objets cibles sont les routes et les zones urbaines. L'auteur a précisé la nécessité d'impliquer des relations spatiales sans les utiliser.

Les auteurs de [43] ont développé un système pour détecter les objets artificiels dans des environnements naturels. L'idée principale est que les objets artificiels ont des surfaces plus uniformes que les objets naturels. Pourtant cette méthode suit une approche de bas niveau fondée sur l'uniformité des textures de l'image [43].

Les auteurs de [24] ont proposé une approche géométrique. Un pont est modélisé comme deux régions rectangulaires placées dans une certaine relation spatiale. Dans chacun des images, on détecte les segments parallèles et on détecte les formes rectangulaires à partir des parallèles. On compare finalement chaque paire de rectangle avec le modèle du pont défini par des règles manuelles pour détecter les ponts. L'originalité de cette approche est l'utilisation des primitives géométriques pour la reconnaissance, en revanche l'application est limitée à la reconnaissance des ponts [24].

III.11 Conclusion

Dans ce chapitre nous avons présenté les approches et les méthodes principales de reconnaissance d'objets en général qui nécessite une chaîne de traitement (bas niveau ; haut niveau) qui se constitue de plusieurs étapes. Nous avons décrit les différents principaux domaines d'application et les défis majeurs de la reconnaissance d'objets dans les images satellitaires.

Dans le chapitre suivant nous allons traiter une séquence vidéo pour détecter des véhicules et la reconnaissance d'objet au niveau haut. Ensuite nous allons traiter des images prises de la vidéo en appliquant les différentes méthodes de prétraitements d'image.

CHAPITRE IV :

Description de l'application

IV.1 Introduction

Le domaine du traitement d'image est composé de plusieurs techniques qui visent à élaborer et analyser des images. Ces techniques sont entièrement dédiées à l'analyse et au traitement d'images.

L'objectif de ce chapitre est de présenter l'application de reconnaissance que nous avons développée sous Matlab qui a pour but d'analyser l'évolution d'une séquence vidéo pour faire la reconnaissance et la détection des objets.

IV.2 Description de l'application

L'application que nous avons développée est composée de deux interfaces

IV.2.1 Interface principale :

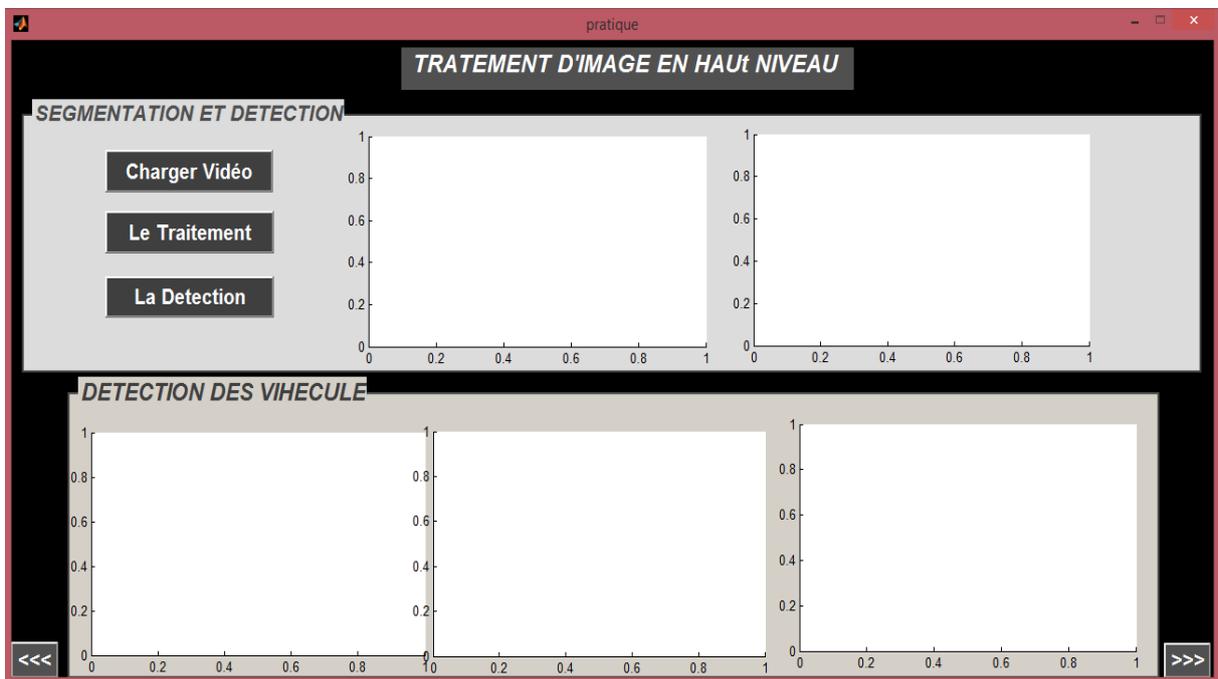


Figure IV.1 : Interface de reconnaissance d'objets et détection des véhicules

Cette interface contient cinq figures pour afficher les images avec leurs résultats après les traitements.

Elle contient aussi trois boutons ; 1^{er} Bouton pour charger la vidéo, 2^{eme} Bouton pour traiter une image obtenu de la vidéo, 3^{eme} Bouton pour la détection des véhicules.

IV.2.2 Interface secondaire

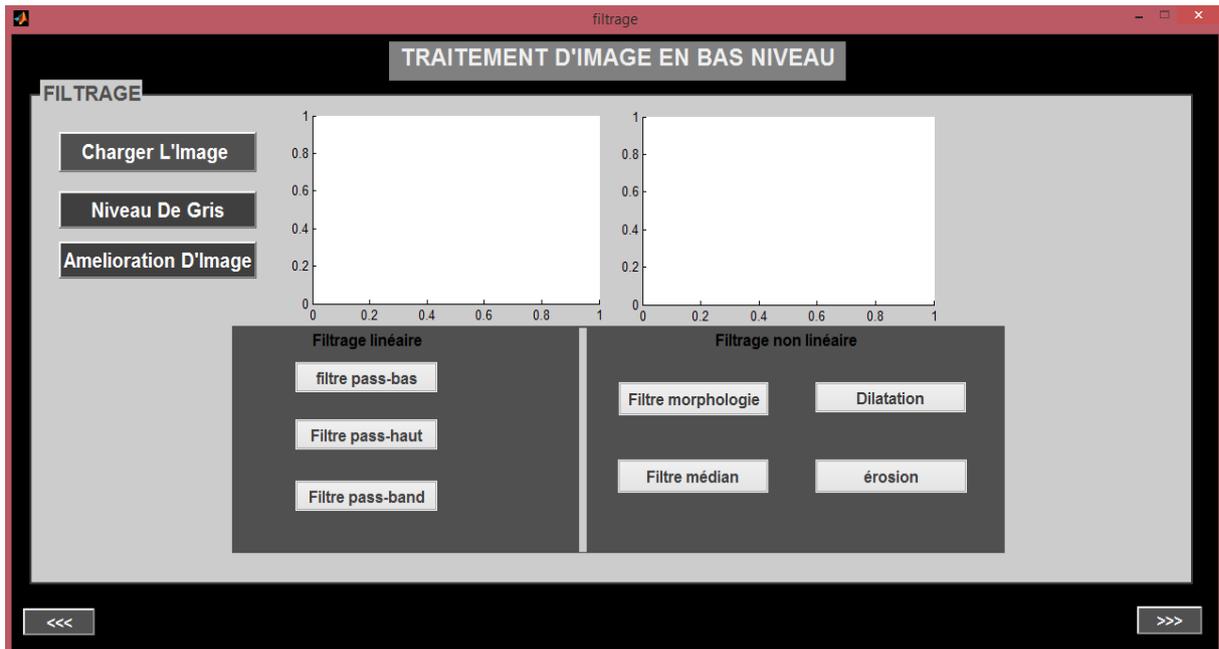


Figure IV.2 interface de traitement d'image en niveau bas

Cette interface contient deux figures pour afficher les images avec leurs résultats après les traitements.

Elle contient aussi 10 boutons : 1^{er} Bouton pour charger l'image, 2^{ème} Bouton pour transformer l'image en niveau de gris, 3^{ème} Bouton pour l'amélioration de la qualité d'image et le reste pour les différentes méthodes de filtrage d'image.

IV.3 Application :

La détection et le comptage de véhicules peuvent être utilisés pour analyser les profils de trafic.

Notre projet montre comment utiliser l'analyse de détection pour détecter et compter les véhicules dans une séquence vidéo, on suppose que la caméra est stationnaire et on va appliquer les traitements de détection d'objets. Plutôt que de traiter immédiatement toute la vidéo, nous avons commencé par l'obtention d'une image vidéo initiale dans laquelle les objets en mouvement sont segmentés en arrière plan.

D'après la figure IV.3, le programme va prendre une image de la vidéo que nous avons choisie pour la traiter :



Figure IV.3 : image prise de la vidéo

a. Segmentation :

Dans cette étape l'image sera convertis en binaire (noir et blanc) pour mieux détecté l'objet cherché, le résultat de cette opération est une image noir et blanc dans laquelle les véhicules sont en blanc et tout le reste est en noir avec la présente de quelques trous en blanc (Figure IV.4).

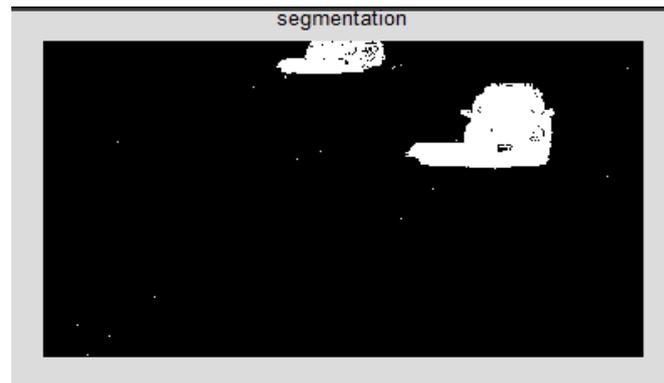


Figure IV.4: segmentation

b. Minimisations des bruits:

La segmentation du premier plan n'est pas parfaite et comporte souvent des trous blanc indésirables. Pour cela on utilise le filtre d'ouverture morphologique (érosion, dilatation) pour éliminer le bruit et pour combler les lacunes dans les objets détectés.

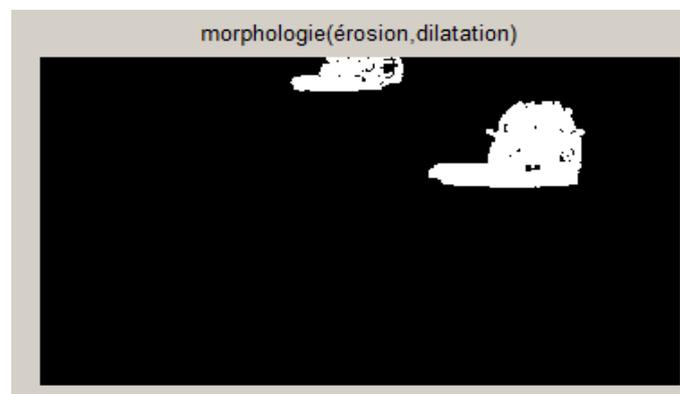


Figure IV.5 : éliminé les pixels isolés

c. Détection de formes:

Une forme peut être prise comme le contour d'un objet. La détection de forme est équivalente à détecter la surface de véhicule et pour mettre en évidence les véhicules détectés, nous dessinons des cases vertes autour d'elles.

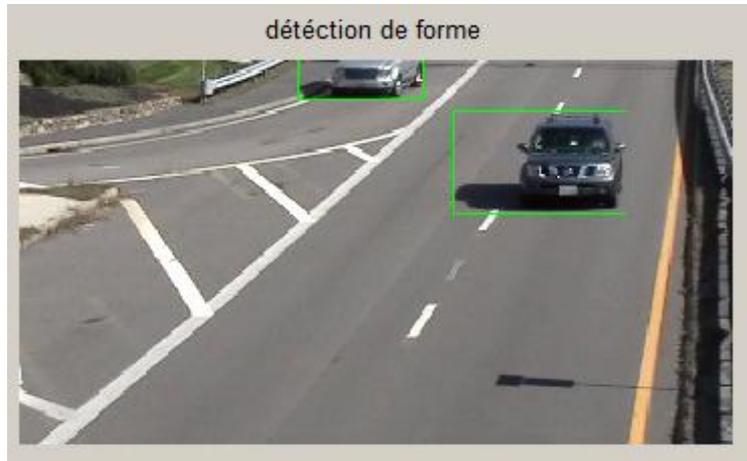


Figure IV.6 : détection de forme

d. Détection et comptage de véhicules :

Dans l'étape finale, nous avons traité les images vidéo restantes qui nous affichent les cases délimitant les véhicules en vertes. Il affiche également le nombre des véhicules détecté prise par une séquence vidéo, tant que ce nombre dépasse 2 (deux véhicules) le programme déclare un encombrement qui nous affiche les cases délimitent en rouge.

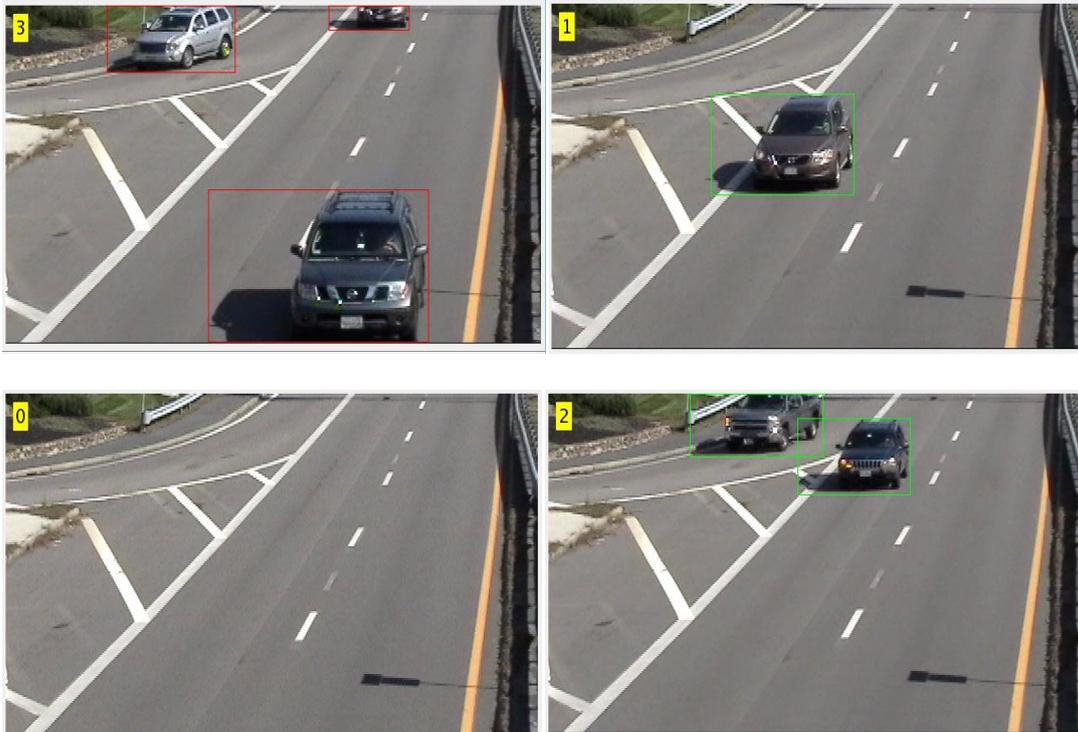


Figure IV.7: détection des véhicules

IV.4 Quelques opérations sur les images pour niveau bas :

a) Chargement des images

Tout d'abord, pour lancer notre application, l'utilisateur doit cliquer sur le bouton « charger l'image », pour charger une image à partir d'une séquence vidéo.

Après avoir chargé une image, on applique les différentes opérations de traitement d'image.



Figure IV.8 : chargement d'image de vidéo

b) Images en niveau de gris :

Cette étape est très importante pour notre application, parce que les images que nous avons choisies sont en couleur. Donc pour faciliter le traitement nous avons converti les images en niveau de gris.

c) L'histogramme :

C'est le nombre des pixels en fonction de niveau de gris. Les résultats obtenus pour ces deux étapes sont présentés dans la figure suivante:

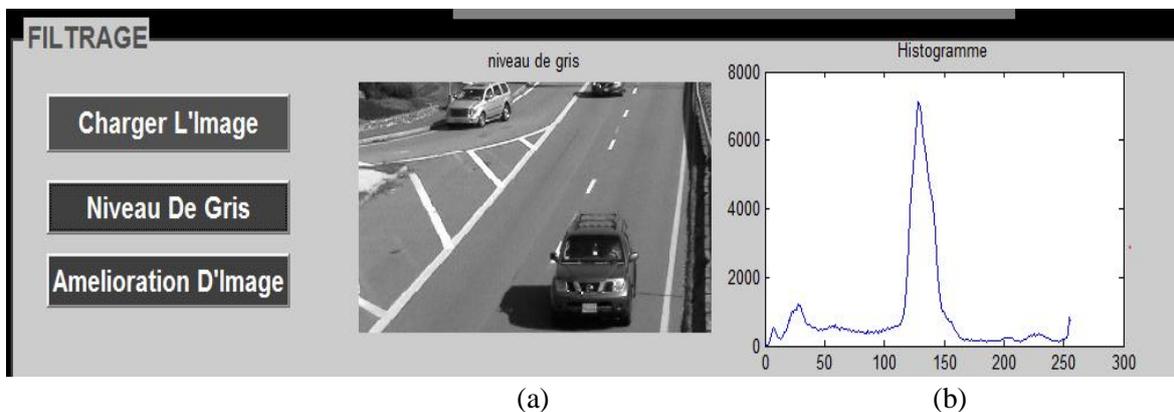


Figure IV.9 : a) niveau de gris, b) l'histogramme

d) Amélioration de la qualité d'image :

Nous avons utilisé cette étape pour améliorer la qualité d'image en appliquant « Amélioration d'image » c'est une dilatation de l'histogramme, c'est à-dire que l'histogramme doit être plus large (doit occuper tout l'intervalle [0 255])

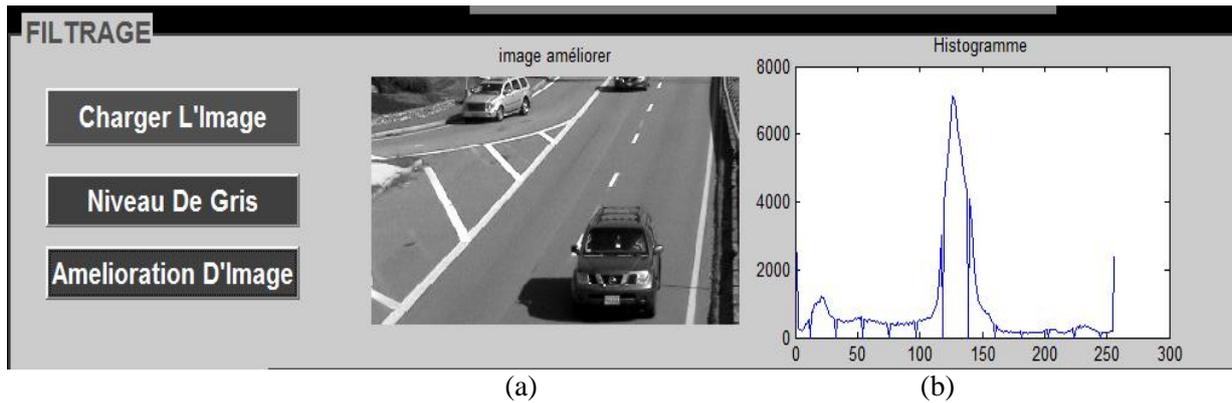


Figure IV.10 : a) image améliorer, b) dilatation de l'histogramme

1) Filtres linéaires

- **Le filtre passe-bas**

Dans ce type des filtres, toutes les fréquences qui dépassent une fréquence bien déterminée seront automatiquement éliminées, le résultat est présenté sur la figure IV.11 :

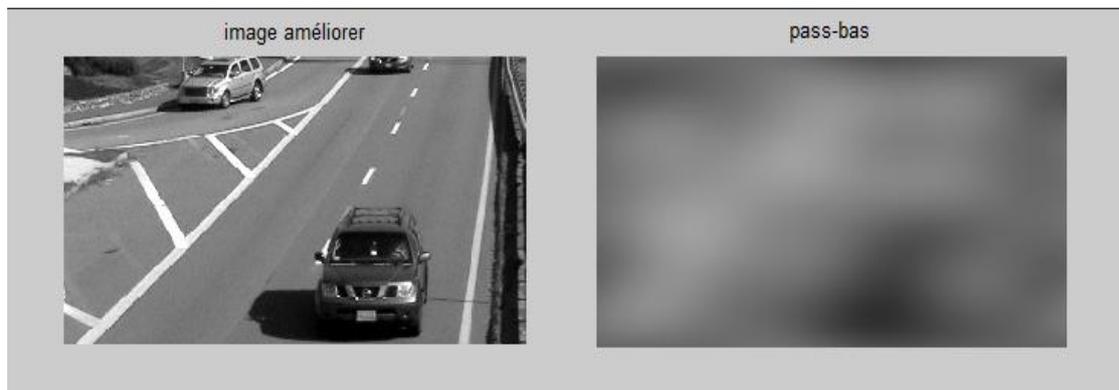


Figure IV.11: filtre passe-bas

- **Le filtre passe-haut**

Les filtres passe-haut éliminent toutes les fréquences qui sont au-dessous de la fréquence de seuil et acceptent les autres fréquences.



Figure IV.12: filtre passe-haut

- **le filtre passe bande**

Ce filtre est un mélange de deux filtres (passe bas, passe haut), donc le filtre éliminent tous les fréquences qui au-dessus a une fréquence max et au-dessous a la fréquence min.

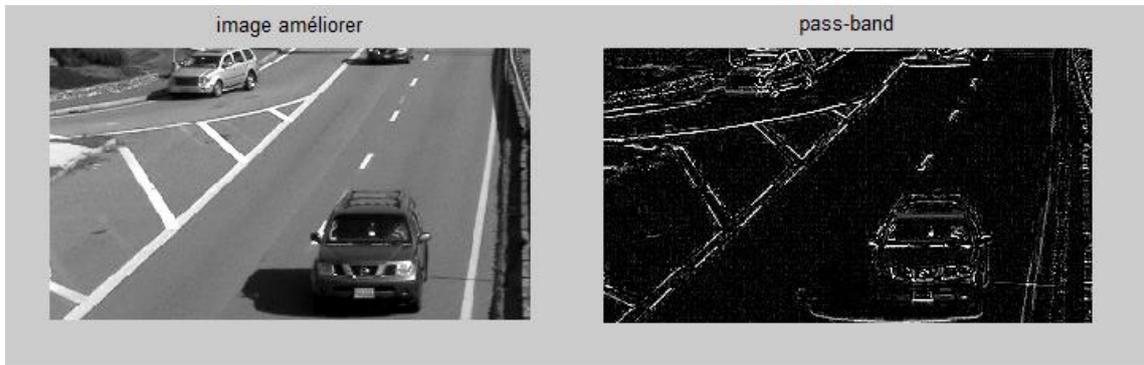


Figure IV.13: filtre passe-bande

2) Filtres non-linéaires

- **Les filtres morphologiques**

Toutes les transformations morphologiques sont définies à l'aide d'un élément structurant et un opérateur. L'élément structurant est un ensemble de pixels



Figure IV.14: filtre morphologie

- **La Dilatation :**

Elle permet d'éliminer les points noirs isolés de l'image. Elle élargit les «pics» et comble les «vallées». La dilatation peut être réitérée jusqu'à ce que l'image ait une valeur constante (la valeur maximale).



Figure IV.15: filtre dilatation

- **L'Érosion :**

Elle permet d'éliminer les points blancs isolés de l'image. Elle abaisse les «pics» et élargit les «vallées». L'érosion peut être réitérée jusqu'à ce que l'image ait une valeur constante (la valeur minimale).



Figure IV.16: filtre érosion

- **Le filtre médian :**

Il supprime le bruit dans l'image. Son algorithme est relativement simple, le filtre remplace la valeur de chaque pixel par la valeur médiane du pixel et de ses voisins.

Nous avons appliqué ce filtrage dans la figure IV.17:



Figure IV.17: filtre médian

IV.5 Conclusion

Durant ce projet, nous avons fait une application pour la reconnaissance d'objets et la détection des véhicules à partir d'une séquence vidéo prise par une caméra de surveillance. Comme dernier étapes, nous avons implémenté quelques techniques de traitement d'images. Pour cela nous avons appliqué plusieurs méthodes (changer l'image en niveau de gris, amélioration, histogramme, filtrage...).

Conclusion générale

L'image satellite est une avancée technologique extraordinaire. Elle est possible grâce au satellite et à leurs envois dans l'espace pour mieux connaître ou même de découvrir autrement notre planète.

Ces imageries permettent de voir nos erreurs que l'être humain commet et aussi les transformations que l'homme donne à la planète. La vie de l'Homme sur terre a complètement changé et la vision de la terre est indéniablement transformée car la planète paraît plus petite et plus accessible grâce à l'imagerie satellitaire.

Au court de ce projet, Nous avons pu retracer :

- ✓ les moyens modernes de suivis des objets en mouvement à travers la première partie le traitement et l'analyse d'image dans la deuxième partie.
- ✓ troisième partie nous avons présenté les méthodes principales de reconnaissance d'objets dans les images.
- ✓ nous avons appliqué quelques différentes méthodes de traitement sur une image prise de la vidéo.
- ✓ pour finir nous avons développé une application sous matlab pour détecter et compter les véhicules dans une séquence vidéo.

Perspectives

Comme perspectives nous proposons :

- Tester notre application sur une caméra stationnaire.
- Développer la solution proposée en utilisant un langage de programmation évolué (c++).
- Utilisation de vidéo haute résolution (HD).
- Evaluer l'application pour l'utiliser comme un radar qui permet de détecter la vitesse des véhicules et leurs matricule.
- Comptage globale des véhicules dans toute la séquence vidéo pour détecter les vrais encombrements.
- Elimination des véhicules redondantes qui sont détectés plus d'une fois.

Bibliographie

- [1] GUEDAT C et BRETTE O ; Systèmes de Géolocalisation, INSA Lyon Département TELECOM SERVICES, 2010.
- [2] Françoise D ; GPS localisation et navigation par satellites, 2^e édition revue et augmentée, Hermes, 2005.
- [3] géolocalisation' USTL mémoire de Master pro télécom Université des Sciences et technologies de Lille, 2007.
- [4] [www.slideserve.fr / présentation-géolocalisation](http://www.slideserve.fr/présentation-géolocalisation), date de consultation mai 2016.
- [5] Guillaume S ; Introduction à la Géolocalisation 802.11, mémoire de master Université Louis Pasteur LSIIT.2008
- [6] Wang X ; Mobile Ad-Hoc Networks : Applications chapters published January 30, 2011.
- [7] Notions fondamentales de télédétection. «<http://www.rncan.gc.ca/sciences-terre/geomatique/imagerie-satellitaire-photos-aeriennes/imagerie-satellitaire-produits/ressources-educatives/9364> » Date de la dernière modification : 2016-08-17.
- [8] DAMIEN L ; Analyse rétro spectrice par télédétection de l'évolution du dépérissement des Sapins pectinés (abies alba) sur le Mont Ventoux Mémoire de master, Ecole Supérieure des Géomètres et Topographes paris, 2010.
- [9] GUYOT G Mémoire d'Ingénieur de l'ESGT 1995 paris.
- [10] Kergomard C ; « LA TÉLÉDÉTECTION AÉRO-SPATIALE » cours Ecole Normale Supérieure Paris.2002
- [11] AKROUR N et CHABI L ; Une plateforme évolutive pour le développement d'approches biomimétiques en segmentation d'image. Mémoire de Master université d'E.S.I, Algérie, 2015.
- [12] BENAMROUZ S et KETTANE S ; Segmentation d'image par les méthodes adaptives basées sur les matrices de cooccurrences. Mémoire de Master université d'UMMTO, 2009.
- [13] BEN YALLOUL A et BEN SALEM S ; Méthodes biomimétiques pour la segmentation d'image. Mémoire Master université d'I.N.I, Algérie, 2007.
- [14] SLIME S ; Environnement de segmentation d'image à base d'une approche biomimétique. Mémoire de Master université E.N.I, Algérie, 2008.
- [15] LAKHDARI M ; Segmentation d'images par contour actif en appliquant les algorithmes génétiques . memoire de Master université I.N.I, Algérie, 2008.
- [16] Lortic B et Lortic C ; L'imagerie radar. Villes et Citadins du Tiers-Monde Paris , 2002
- [17] HOUASSINE C ; Segmentation d'images par une approche biomimétique hybride, mémoire de Magister université Boumerdes, Algérie, 2012.
- [18] LECOEUR J et BARILLOT C ; Segmentation d'images cérébrales : État de l'art. IRISA, mémoire de master, Rennes France. Version 3, 2008.
-

Bibliographie

- [19] Cerella J; Pigeons and perceptrons. Journal Pattern Recognition 1986.
- [20] Venkateswar V et Chellappa R ; Extraction of straight lines in aerial images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992.
- [21] Auclair A, Cohen L et Vincent N ; How to use SIFT vectors to analyze an image with database templates. 5ème Atelier international sur la récupération multimédia adaptative, Paris, France, Juillet 5-6 2007.
- [22] Harris C et Stephens M; A Combined Corner and Edge Detector, la 4^{ème} Conférence de Vision ALVEY, 1988.
- [23] Schmid C et Mohr R; Local gray value invariants for image retrieval. IEEE Trans. Pattern Anal. Machine Intel, 1997.
- [24] Loménie N, Trias-Sanz R et Barbeau J ; Integrating textural and geometric information for an automatic bridge detection. conference, Toulouse, France ,2003.
- [25] Lindeberg T; Feature detection with automatic scale selection. International Journal. Computing,1998.
- [26] Perona P et Malik J ; Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell, 1990.
- [27] Canny J ; A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell, 1986.
- [28] Andrew A; Méthodes définies par le niveau et méthodes de marche rapide: Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science,2000.
- [29] Shotton J, Blake A et Cipolia R; Multi-scale categorical object recognition using contour fragments. journal IEEE Transactions on Pattern Analysis and Machine Intelligence 2007.
- [30] Ferrari V, Tuytelaars T et Gool L ; Object detection by contour segment networks,2006.
- [31] Duda R et Hart P; Use of the Hough transformation to detect lines and curves in pictures, 1972.
- [32] Venkateswar V et Chellappa R ; Extraction of straight lines in aerial images. Journal IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [33] Desolneux A, Moisan L. et Morel J; Meaningful alignments. Int. J. Comput, 2000.
- [34] Garnesson P et Giraudon, G ; L'approximation polygonale : bilans et perspectives. Rapport technique RR-1621, INRIA - Sophia Antipolis, Equipe : PASTIS, 1992.

Bibliographie

- [35] Iqbal Q et Aggarwal J; Retrieval by classification of images containing large manmade objects using perceptual grouping. *Pattern Recognition Journal*, 2002.
- [36] Kavallieratou E, Fakotakis N et Kokkinakis G; Handwritten character recognition based on structural characteristics, paris 2002.
- [37] Khotanzad A et Hong Y; Invariant image recognition by Zernike Moments. *IEEE Trans. Pattern Anal. Mach. Intell*,1990.
- [38] Crevier D et Lepage R ; Knowledge-based image understanding systems a survey. *Computer. Vision. and Image Underst*, 1997.
- [39] Turk M et Pentland A ; Face recognition using eigenfaces, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*,1991.
- [40] Lorette A, Descombes X et Zerubia J ; Texture analysis through a markovian modelling and fuzzy classification : Application to urban area extraction from satellite images. *International Journal of Computer Vision*, 2000.
- [41] Bordes G ; Interprétation d'images satellitaires à l'aide de données externes. Thèse de doctorat, Université Marne-la-Vallée, 1997.
- [42] Forssen P ; Detection of man-made objects in satellite images. Mémoire de D.E.A., Department of Electrical Engineering, Linkopings University, Sweden, 1997.
- [43] Caron Y, Makris P et Vincent N; A method for detecting artificial objects in natural environments. In *ICPR 02: Actes de la 16^{eme} Conférence internationale sur la reconnaissance*, Washington, DC, USA. IEEE computer Society, 2002.

ANNEXE

Interface principale

```
function pushbutton9_Callback(hObject, eventdata, handles)
foregroundDetector = vision.ForegroundDetector('NumGaussians', 4,
...
    'NumTrainingFrames', 50);

videoReader = vision.VideoFileReader('visiontraffic.avi');

for i = 1:140
    frame = step(videoReader);
    foreground = step(foregroundDetector, frame);
end
se = strel('square', 3);
filteredForeground = imopen(foreground, se);

blobAnalysis = vision.BlobAnalysis('BoundingBoxOutputPort', true,
...
    'AreaOutputPort', false, 'CentroidOutputPort', false, ...
    'MinimumBlobArea', 150);
bbox = step(blobAnalysis, filteredForeground);

result = insertShape(frame, 'Rectangle', bbox, 'Color', 'green');

numCars = size(bbox, 1);
result = insertText(result, [10 10], numCars, 'BoxOpacity', 1, ...
    'FontSize', 14);

videoPlayer = vision.VideoPlayer('Name', 'détection des véhicules');
videoPlayer.Position(3:4) = [650,400]; %la fenetre
se = strel('square', 3); %le filtre morphologique

while ~isDone(videoReader)

    frame = step(videoReader); % lir la suite des séquences images
    détecté

    foreground = step(foregroundDetector, frame);

    %on applique le filtre morphologique(erosion,dilatation)
    filteredForeground = imopen(foreground, se);

    bbox = step(blobAnalysis, filteredForeground);

    %cases vertes d'élimittant les véhicules
    result = insertShape(frame, 'Rectangle', bbox, 'Color',
'green');

    %compter le nombre des véhicules
    for i=0

        numCars = size(bbox, 1);
        numCars = numCars+i;
```

ANNEXE

```
    result = insertText(result, [10 10], numCars, 'BoxOpacity', 1,
...
    'FontSize',20);
    if (numCars >= 3)
        result = insertShape(frame, 'Rectangle', bbox, 'Color',
'red');
        result = insertText(result, [10 10], numCars, 'BoxOpacity',
1, ...
    'FontSize',20);
    end

    end
step(videoPlayer, result); title('detection des véhicules') %
afficher le résultat
end

release(videoReader); % quitter le vidéo
```

Interface secondaire

```
function pushbutton1_Callback(hObject, eventdata, handles)
nom=uigetfile('*.*.');
I=imread(nom);
axes(handles.axes4);
axes(handles.axes4);
    imshow(I, 'Parent', handles.axes4); title('image prise de vidéo')
handles.I=I;
guidata(hObject, handles);

function pushbutton2_Callback(hObject, eventdata, handles)
%filtre pass-band
I=handles.I;
J=rgb2gray(I);
J=double(J)/255.0;
h=-ones(3,3); % Definition de h
h(2,2)=8;
F=filter2(h,J);
imshow(F, 'Parent', handles.axes5); title('pass-band');
handles.I=I;
guidata(hObject, handles);

function pushbutton3_Callback(hObject, eventdata, handles)
I=handles.I;
i=rgb2gray(I);
j=imnoise(i, 'salt & pepper', 0.02);
axes(handles.axes4);
imshow(j, 'parent', handles.axes4); title('image bruité');
s= strel('disk',1);
k=imopen(j,s);
n=imclose(k,s);
p=imopen(n,s);
axes(handles.axes5);
imshow(p, 'parent', handles.axes5); title('filtre median');
```

ANNEXE

```
function pushbutton4_Callback(hObject, eventdata, handles)
I=handles.I;
J=rgb2gray(I);
s= strel('square',15) ;
i=imerode(J,s);
j=imdilate(i,s);
imshow(j,'parent',handles.axes5);title('morphologique');

function pushbutton5_Callback(hObject, eventdata, handles)
I= handles.I;
Ic=rgb2gray(I);
SE = strel('ball',5,5); % définition de l'élément structurant
Idilat = imdilate(Ic,SE);
imshow(Idilat,'parent',handles.axes5);title('dilatation');

function pushbutton6_Callback(~, eventdata, handles)
%filtre pass-haut
I=handles.I;
F=fftshift(fft2(I));
%calcul de la taille de l'image;
M=size(F,1);
N=size(F,2);
P=size(F,3);
H1=ones(M,N);
D0=2;
M2=round(M/2);
N2=round(N/2);
H1(M2-D0:M2+D0,N2-D0:N2+D0)=0;
for i=1:M
for j=1:N
G(i,j)=F(i,j)*H1(i,j);
end
end
g=ifft2(G);
imshow(255-abs(g), [0,255], 'Parent',handles.axes5);title('pass-
haut');

function pushbutton7_Callback(hObject, ~, handles)
I=handles.I;
F=fftshift(fft2(I));
M=size(F,1);
N=size(F,2);
P=size(F,3);
H0=zeros(M,N);
D0=2;
M2=round(M/2);
N2=round(N/2);
H0(M2-D0:M2+D0,N2-D0:N2+D0)=1;
for i=1:M
for j=1:N
G(i,j)=F(i,j)*H0(i,j);
end
end
g=ifft2(G);
imshow(abs(g), [0,255], 'Parent',handles.axes5);title('pass-bas');
```

ANNEXE

```
function pushbutton10_Callback(hObject, eventdata, handles)
I= handles.I;
Ic=rgb2gray(I);
SE = strel('ball',5,5); % définition de l'élément structurant
Ierod = imerode(Ic,SE);
imshow(Ierod, 'parent',handles.axes5);title('érosion');
```

```
function pushbutton12_Callback(hObject, eventdata, handles)
I=handles.I;
I1=rgb2gray(I);
axes(handles.axes4);
imshow(I1, 'parent',handles.axes4);title('niveau de gris');
h=imhist(I1);
axes(handles.axes5);
plot(handles.axes5,h);title('Histogramme');
handles.I1=I1;
guidata(hObject,handles);
```

```
function pushbutton13_Callback(hObject, eventdata, handles)
I1=handles.I1;
u=imadjust(I1);
axes(handles.axes4);
h=imshow(u, 'parent',handles.axes4);title('image améliorer');
j=imhist(u);
axes(handles.axes5);
plot(handles.axes5,j);title('Histogramme');
```