

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE

UNIVERSITÉ ABDELHAMID IBN BADIS MOSTAGANEM



**Faculté des Sciences Exactes et d'Informatique**  
**Département de Mathématiques et d'Informatique**  
**Filière Informatique**

*Mémoire de fin d'étude pour l'obtention du diplôme de*  
*Master II en informatique*  
**Spécialité: Systèmes d'Information Géographiques**

## **Thème**

**ETUDE D'ALGORITHMES DE DÉTERMINATION DE NOMBRE DE  
CLUSTERS A PRIORI : Application aux Données satellitaires**

**Présenté par :**

- Hedjadj Aboubaker
- Kadour Ahmed

**Devant le jury composé de:**

- KAID SLIMANE et BESNASSI : examinateur
- K Labed et H Mahi : encadreur

**Année Universitaire 2014/ 2015**



## SOMMAIRE

Résume.....	3
Introduction générale .....	4
La télédétection et L'imagerie .....	5
I.1.    Introduction .....	6
I.2.    Télédétection .....	6
I.2.a Définition.....	6
I.2.b Principe de base .....	6
I.2.c Quels sont les différents systèmes de télédétection ? .....	7
I.3.    Les bases physiques de la télédétection.....	7
I.3.a Le Rayonnement électromagnétique .....	7
I.3.b Le spectre électromagnétique .....	9
I.4.    Les éléments de la Télédétection.....	9
I.4.b Objets observés.....	11
I.4.c Les capteurs .....	11
I.5.    Imagerie .....	12
I.5.a Définition d'image.....	12
I.5.b Image numérique .....	12
I.5.c Caractéristiques d'une image numérique .....	12
I.5.d Traitement d'images numérique .....	13
I.5.e Techniques d'extraction de l'information.....	13
I.6.    Domaines d'application [4] .....	14
I.6.a Agriculture.....	14
I.6.b Applications maritimes.....	14
I.6.c Cartographie 2D ou 3D.....	15
I.6.d Risques naturels et industriels .....	15
I.7.    Conclusion .....	15
Classification et l'algorithme de VAT .....	16
II.1.   Introduction .....	17
II.2.   Classification .....	17
II.2.a Définition .....	17
Distance euclidien : .....	17
Une Classe (cluster).....	18

II.2.d.	Principe des algorithmes classification.....	18
II.2.c	Méthodes de classification.....	18
II.3.	Tendance à évaluation visuelle VAT (visual assument tendency).....	21
II.3.a.	Définition.....	21
II.3.b.	Matrice de dissemblance et image de VAT.....	22
II.3.c.	Algorithme De VAT.....	23
	Exemple.....	23
II.3.d.	Avantage et les limite de l'outil le VAT.....	24
II.4.	Version Amélioré VAT (improved VAT (iVAT).....	25
II.4.a	Définition.....	25
II.4.b	Algorithme de IVAT.....	25
II.5.	Extraction de nombre de cluster (Cluster Count Extract (CCE)).....	26
II.5.a	Définition.....	26
IV.5.b	Algorithme.....	26
IV.6.	Méthode L'extraction de bloc foncé (Dark Block Extraction(DBE)).....	27
IV.6.a	Définition.....	27
IV.6.b	Algorithme DBE.....	27
II.7.	Conclusion.....	27
Conception ET implementation.....		28
III.1.	Introduction.....	29
III.2.	Ressources Utilisées.....	29
III.2.a.	Les ressources matérielles :.....	29
III.2.b.	Les ressources logicielles :.....	29
III.3.	Conception.....	29
III.4.	Données utilisées.....	30
III.5.	Présentation de l'application.....	32
III.5.a.	Application d'algorithme VAT.....	34
III.5.b.	Application d'algorithme iVAT.....	35
III.6.	Discussions.....	37
III.7.	Conclusion.....	38
Conclusion générale.....		39
Bibliographie.....		40

## Résumé

Le clustering est une technique de classification non supervisée dont l'objectif est d'organiser un ensemble de données non labialisées en cluster, de telle manière que les données du même cluster aient une valeur de similarité alors qu'à l'opposé, deux données de deux cluster différents aient une grande valeur de dissimilarité. En traitement d'images satellitaires les algorithmes de classification non supervisée les plus populaires sont : l'algorithme K-means et l'algorithme ISODATA, cependant l'inconvénient de ces algorithmes est leur sensibilité aux conditions initiales : d'une part, le choix du nombre de cluster  $C$  qui a une grande influence sur la qualité du résultat final, et d'autre part, le choix aléatoire de centroïde qui conditionne la partition finale.

Face à ce constat, dans ce travail on a étudié et implémenté des algorithmes visant à déterminer le nombre de cluster a priori en exploitant une matrice dite de dissimilarité. Ces algorithmes sont connus sous le nom de **visuel Assessment Tendency (VAT)** et comprennent plusieurs variantes

: VAT, iVAT, etc.

L'objectif de ce travail est d'utiliser et de tester quelques variantes de ces algorithmes à la fois sur des données synthétiques et des données réelles (image satellitaire) afin de sélectionner le plus performant d'entre eux.

**Mot clés :** Clustering, K-means, VAT, iVAT, Matrice dissimilarité, image satellitaire.

## Introduction générale

L'être humain, poussé par sa curiosité, essayait d'expliquer tout ce qu'il voit en utilisant ses deux yeux ainsi que son cerveau pour capter et traiter les différentes informations relatives aux objets de son environnement système a été le premier processus de la télédétection.

Dans l'époque actuelle, nous avons remplacé les yeux avec des satellites et le cerveau avec des ordinateurs, qui produisent des photos grand et variété, Ces diverses doivent avoir un processus d'extraction d'informations à travers plusieurs étapes, la classification est la plus importante étape.

Le but le plus simple de la classification est de regrouper en classes homogènes un ensemble d'individus. Et nous utilisons des algorithmes pour faire le regroupement des algorithmes supervise (K-plus proches voisins) et des algorithmes non supervise (k-means), Mais tous les algorithmes de classification ont un gros problème ils ne peuvent pas calculer le nombre de divisions donc doit toujours, être saisi par l'utilisateur comme un entrée.

Pour résoudre le problème de classification ils ont programmé quelques algorithmes. Le plus important algorithme est le VAT.

L'algorithme de VAT est une méthode déterminé le nombre possible de clusters dans un ensemble d'objets visuellement.

Malgré l'évolution de cette algorithme, Mais il est confronté beaucoup de problèmes Le plus important Image extraite avec VAT parfois n'être pas clair en raison de la complexité de l'image ou de la taille, Ils ont développé une fonction pour résoudre ce problème appelé VAT amélioré (IVAT) C'est le premier problème, le second L'algorithme encore besoin de soutien humain pour compter le nombre de cases noires dans l'image de VAT Il produit un algorithme de prise en charge appelé CCE cluster compte d'extraction .

Après résolu ces problèmes l'ordinateur peut diviser presque n'importe quelle ensembles de données ou photo sans intervention extérieure. Cela a donné un coup de pouce pour de nombreux domaines comme la télédétection et exploration de l'espace et les robotiques. la recherche et gestion de web.

# **CHAPITRE 1**

**La télédétection et L'imagerie**

## I.1. Introduction

Le terme télédétection désigne, dans le concept général, un ensemble de techniques permettant de mesurer à distance (c'est-à-dire sans contact) des grandeurs physiques caractéristiques des objets ou des phénomènes étudiés. Toutefois, en pratique, il est surtout utilisé pour les applications spatiales et aéroportées, et concerne essentiellement les techniques mises en œuvre pour l'observation de la surface de la terre et d'autres planètes. Ces systèmes nécessitent l'utilisation de capteurs qui sont embarqués sur des ballons, des avions, des navettes ou des satellites.

## I.2. Télédétection

### I.2.a Définition

C'est l'ensemble des connaissances et techniques utilisées pour déterminer des caractéristiques physiques et biologiques d'objets par des mesures effectuées à distance, sans contact matériel avec ceux-ci.

C'est l'ensemble des connaissances et techniques nécessaires pour interpréter divers « objets » par leurs comportements spectraux et leurs distributions spatiales à l'aide de mesures spécifiques effectuées à distance [1].

### I.2.b Principe de base

#### La source d'énergie

Est l'élément qui "éclaire" la cible en émettant une onde électromagnétique (flux de photons).

#### La cible

Est la portion de la surface terrestre observée par le vecteur.

#### Le vecteur

Ou plate-forme de télédétection mesure l'énergie solaire (rayonnement électromagnétique) réfléchi par la cible. Cela peut être un satellite, un avion, un drone, un hélicoptère, un cerf-volant.

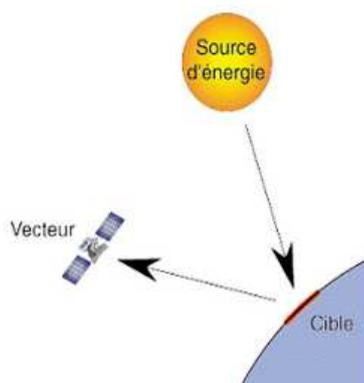


Figure 1 Schéma expliquant le principe de la télédétection

### I.2.c Quels sont les différents systèmes de télédétection ?

Si les ondes électromagnétiques mises en jeu dans le processus sont émises par un capteur (exemple : un radar) puis recueillies par ce même capteur après interaction avec la zone terrestre observée, on parle de **télédétection active** (figure 2). Si le capteur (exemple : un radiomètre) recueille directement la lumière visible ou infrarouge émise ou diffusée par la zone terrestre observée, on qualifie les ondes analysées d'ondes électromagnétiques naturelles et on parle de **télédétection passive** (figure 3).

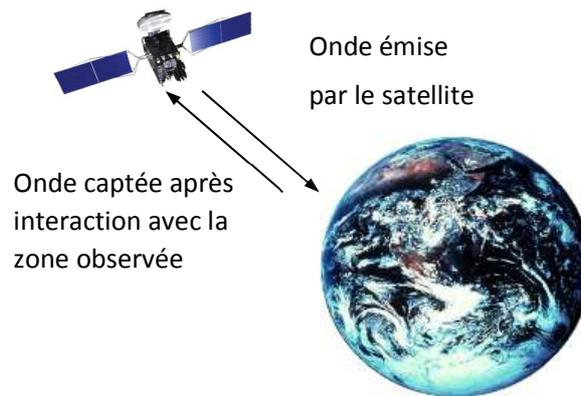


Figure 2 Principe de la télédétection active

### I.3. Les bases physiques de la télédétection

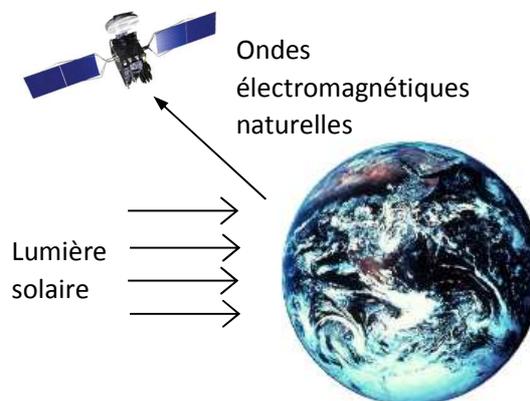


Figure 3 Principe de la télédétection passive

#### I.3.a Le Rayonnement électromagnétique

Premièrement, une source d'énergie sous forme de rayonnement électromagnétique est nécessaire pour illuminer la cible, à moins que la cible ne produise elle-même cette énergie.

Selon la théorie des ondes, tout rayonnement électromagnétique possède des propriétés fondamentales et se comporte de façon prévisible. Le rayonnement électromagnétique est composé d'un champ électrique (E) et d'un champ magnétique (M). Le champ électrique varie en grandeur et est orienté de façon perpendiculaire à la direction de propagation du rayonnement.

Le champ magnétique est orienté de façon perpendiculaire au champ électrique. Les deux champs se déplacent à la vitesse de la lumière. (C) (figure4).

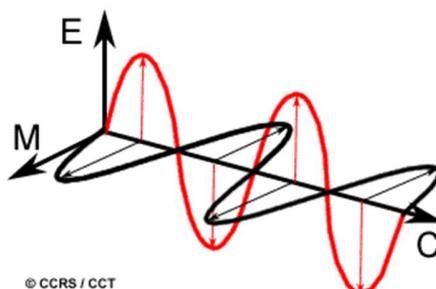


Figure 2 Onde électromagnétique

Pour comprendre la télédétection, il est indispensable de saisir les deux composantes du rayonnement électromagnétique que sont la longueur d'onde et la fréquence (figure 5).

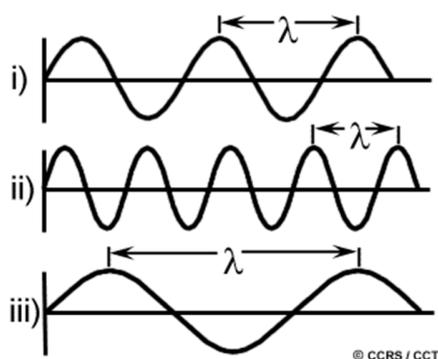


Figure 3 La longueur d'onde

La longueur d'onde équivaut à la longueur d'un cycle d'une onde, ce qui correspond à la distance entre deux crêtes successives d'une onde. La longueur d'onde est représentée habituellement par la lettre grecque lambda ( $\lambda$ ), et est mesurée en mètres ou en l'un de ces sous-multiples tels que les **nanomètres** (nm,  $10^{-9}$  mètre), **micromètres** ( $\mu\text{m}$ ,  $10^{-6}$  mètre) ou **centimètres** (cm,  $10^{-2}$  mètre). La fréquence représente le nombre d'oscillations par unité de temps. La fréquence est normalement mesurée en Hertz (Hz) (c.-à-d. en oscillations par seconde) ou en multiples d'Hertz. La formule suivante illustre la relation entre la longueur d'onde et la fréquence :

$$C = \lambda \nu$$

Ou' :

$\lambda$ =longueur d'onde

$\nu$ = fréquence

C=vitesse de lumière

La longueur d'onde et la fréquence sont donc inversement proportionnelles, c'est-à-dire que plus la longueur d'onde est petite, plus la fréquence est élevée, et plus la longueur d'onde est grande, plus la fréquence est basse. Afin de comprendre l'information tirée des données de télédétection, il est essentiel de bien saisir les caractéristiques du rayonnement électromagnétique. Nous examinerons maintenant la classification du rayonnement électromagnétique.

### I.3.b Le spectre électromagnétique

C'est le résultat de la décomposition du rayonnement électromagnétique en ses fréquences constituantes. Il s'étend des courtes longueurs d'onde (dont font partie les rayons gamma et les rayons X) aux grandes longueurs d'onde (micro-ondes et ondes radio).

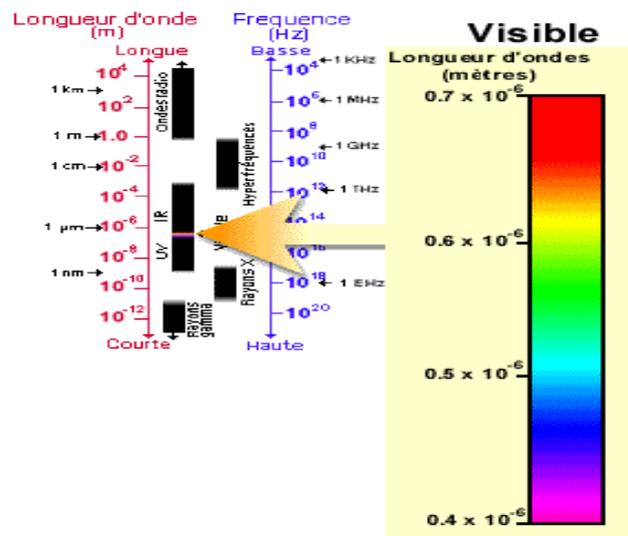


Figure 06 : le spectre électromagnétique.

## I.4. Les éléments de la Télédétection

### I.4.a Les plates-formes

Quel que soit le degré de sophistication des instruments utilisés, en l'absence d'un moyen de quitter la surface de la terre, la télédétection ne permet pas d'acquérir une vision d'ensemble acceptable de notre planète. C'est la raison pour laquelle la télédétection est une technologie relativement nouvelle.

- **Les Avions**

De nos jours, l'une des méthodes les plus évidentes pour "prendre une photo" de la Terre à Distance consiste à faire en s'éloignant de sa surface (en volant, par exemple). Pour pouvoir prendre des photos, les avions sont équipés d'appareils photo. Les avions se caractérisent (du point de vue de la télédétection) par le fait qu'ils volent à une altitude relativement faible (seulement quelques kilomètres au-dessus de la surface) et ne peuvent par conséquent prendre en photo que des portions de territoire limitées, avec de nombreux détails (voitures, personnes, arbres, etc.).

Pour que les avions d'étude puissent voler, les conditions météo doivent être suffisamment bonnes, de manière à ce qu'il soit possible de prendre de nombreuses photographies ; les photos prises par les appareils photos embarqués sur ces avions sont par conséquent généralement assez claires (peu de nuages). Cela étant, les avions ne peuvent pas voler à tout moment (durant la nuit ou par temps de brouillard ou de pluie, alors qu'il est possible que des images soient nécessaires à ces moments).

- **Les satellites**

Les satellites ne volent pas. Ils se déplacent en suivant une orbite. L'orbite est la trajectoire suivie par un corps céleste autour d'un autre corps céleste plus grand. Généralement, l'orbite a une forme quasi-circulaire.

Aujourd'hui, les satellites sont des plates-formes utilisées communément en télédétection ; ils véhiculent des capteurs extrêmement variés, et souvent spécialisés dans l'observation de la météo, des paysages ou catastrophes naturelles, de la végétation,....

Certains capteurs sont même capables de "voir" à travers les nuages ou d'acquérir de l'imagerie de nuit. Deux grands avantages des satellites par rapport aux avions (figure7), c'est qu'ils peuvent prendre des images d'étendues extrêmement vastes et que la même surface peut être observée systématiquement à chaque passage du satellite.



Figure 7 Image prise par un avion



Image satellitaire

Différents types de satellites :

Nous pouvons classifier les satellites suivant leurs missions ou leurs orbites. Suivant les différentes missions des satellites nous pouvons citer :

- Les satellites d'astronomie.
- Les satellites de navigation.
- Les satellites météorologiques.
- Les satellites militaires.
- Les satellites de communication.

#### **I.4.b Objets observés**

La mise en œuvre des techniques de télédétection dépend ce qui doit être observé. Par exemple, les paramètres orbitaux sont étroitement liés aux réquisits de l'observation. Ainsi, La Terre peut être observée à différentes échelles.

#### **I.4.c Les capteurs**

La fonction d'un capteur consiste à détecter le signal radiatif émis ou réfléchi par la surface et à l'enregistrer soit sous forme analogique, soit sous forme numérique. Trois grands types de capteurs peuvent être distingués:

- les appareils photographiques,
- les radiomètres imageurs,
- les capteurs actifs (radars).

## I.5. Imagerie

### I.5.a Définition d'image

Une image est une représentation visuelle, de quelque chose (objet, être vivant). Elle peut être naturelle (ombre, reflet) ou artificielle (peinture, photographie), visuelle ou non, tangible ou conceptuelle (métaphore), elle peut entretenir un rapport de ressemblance directe avec son modèle ou au contraire y être liée par un rapport plus symbolique.

### I.5.b Image numérique

Le terme d'image numérique désigne, dans son sens le plus général, toute image qui a été acquise, traitée et sauvegardée sous une forme codée représentable par des nombres (valeurs numériques).

Une image numérique en niveaux de gris est un tableau de valeurs. Chaque case de ce tableau, qui stocke une valeur, se nomme un pixel. En notant  $n$  le nombre de lignes et  $p$  le nombre de colonnes de l'image, on manipule ainsi un tableau de  $n \times p$  pixels.

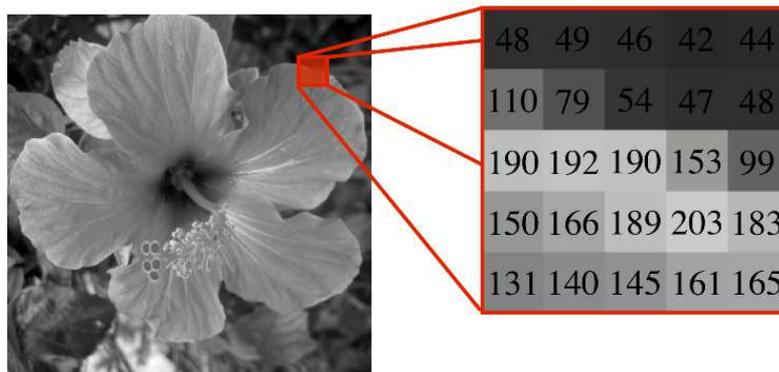


Figure 8 Image numérique

### I.5.c Caractéristiques d'une image numérique

#### Pixel

Le pixel correspond donc à un point de nuance de couleur qui, assemblé à beaucoup d'autres, forme une image dite matricielle.

La quantité d'information que véhicule chaque pixel donne des nuances entre images monochromes et images couleurs. Dans le cas d'une image monochrome, chaque pixel est codé sur un octet, et la taille mémoire nécessaire pour afficher une telle image est directement liée à la taille de l'image.

Dans une image couleur (R.V.B.), un pixel peut être représenté sur trois octets : un octet pour chacune des couleurs : rouge (R), vert (V) et bleu (B).

**Dimension** Le nombre de lignes de matrice (image) multiplié par le nombre de colonnes donne le nombre total de pixels dans une image.

**Résolution** La résolution d'une image composée de points est définie par la densité des points par unité de surface. La résolution permet de définir la finesse de l'image. Plus la résolution est grande, plus la finesse de l'image est grande.

**Bruit** Les images sont parfois de mauvaise qualité. Un exemple typique de défaut est le bruit qui apparaît quand une photo est sous-exposée, c'est-à-dire qu'il n'y a pas assez de luminosité. Ce bruit se manifeste par de petites fluctuations aléatoires des niveaux de gris. La figure ci-dessous montre une image bruitée.

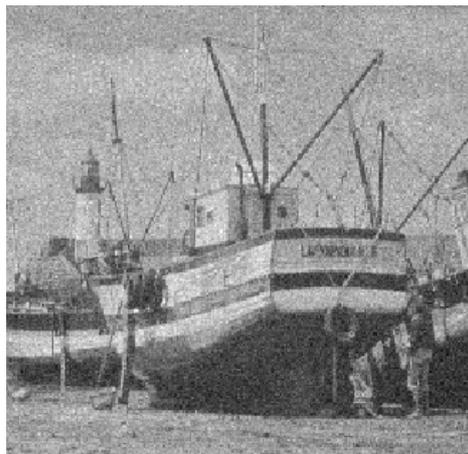


Figure 9 Image bruitée

**Histogramme** Un histogramme est un graphe comptant combien de pixels sont à chaque niveau entre noir et blanc. La hauteur du graphique à représenter chaque point dépend de combien de pixels sont lumineux.

#### **I.5.d Traitement d'images numérique**

Le Traitement d'images numériques est l'utilisation d'algorithmes informatiques pour exécuter des traitements sur des images numériques. Comme une sous-catégorie ou un champ de traitement numérique du signal, le traitement d'image numérique présente de nombreux avantages sur le traitement de l'image analogique. Il permet un éventail beaucoup plus large d'algorithmes pouvant être appliqué aux données d'entrée et peut éviter des problèmes tels que l'accumulation de bruit et distorsion lors du traitement du signal.

##### **La restauration**

Qui a pour but de produire une image la plus proche de la réalité physique de la scène observée.

##### **L'amélioration**

Qui a pour but de satisfaire l'œil de l'observateur humain.

##### **La compression**

Qui a pour but de faciliter le traitement et surtout le stockage des images par une réduction adéquate de leur volume d'information.

#### **I.5.e Techniques d'extraction de l'information**

##### **La segmentation**

Qui a pour but de partitionner l'image en régions homogènes.

### **La reconnaissance de formes**

Qui permet d'associer un objet ou une propriété à certaines régions de l'image.

### **La classification**

Qui permet le partitionnement de l'image en régions homogènes appelées classes, ces dernières sont construites à partir des attributs de pixels (niveau de gris, texture, *etc.*).

## **I.6. Domaines d'application [4]**

### **I.6.a Agriculture**



Figure 10 Irrigation dans la région de Lethbridge, Canada - Spot 5 - 2,5 m

- Classification des types de cultures et de sol
- Evaluation de la santé des cultures
- Estimation de la production totale d'une récolte
- Surveillance de conformité aux lois et traités.

### **I.6.b Applications maritimes**

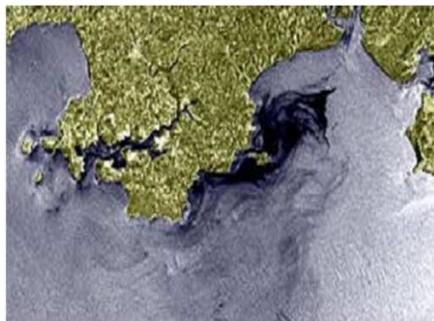


Figure 41 Image Radar SAT des côtes du pays de Galles

- Contrôle des activités de pêche et du trafic maritime
- Localisation des nappes de pollution
- Gestion de l'environnement côtier et océanique

### I.6.c Cartographie 2D ou 3D

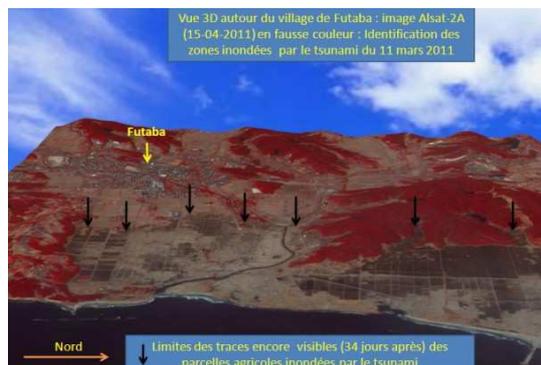


Figure 12 Vue 3d autour du village de futaba (Alsat-2a)

- Créer ou mettre à jour des cartes topographiques ou des plans cadastraux
- Optimiser des campagnes de relevés de terrain
- Renseigner des bases de données cartographiques
- Mettre en place un projet de développement durable

### I.6.d Risques naturels et industriels

- Mettre à jour les plans de prévention des risques
- localiser rapidement les zones affectées et cartographier les dégâts
- améliorer les modèles de prévision et de simulation des phénomènes à risques.

## I.7. Conclusion

Au terme de ce chapitre, nous avons vu un aperçu général sur la télédétection qui est une science vaste, ainsi nous avons survolé les concepts fondamentaux de l'image numérique et les traitements qu'elle peut subir.

La classification d'image devenue une nécessité face à la croissance de la demande d'images et la complexité de celle-ci. Au cours du chapitre suivant, nous présenterons différentes approches de classification applicable sur les images numériques.

# **CHAPITRE 2**

**Classification et l'algorithme de VAT**

## II.1. Introduction

La classification de données situées dans un espace de grande dimension est un problème délicat qui apparaît dans de nombreuses sciences telles que l'analyse d'images, applications web.

La classification est une méthode mathématique d'analyse de données, pour simplifier les données, on les regroupe en classes de telle sorte que les individus d'une même classe soient le plus comparables possible et que les classes soient le plus différents possibles.

Mais avant d'utiliser une méthode de regroupement (densité hiérarchique, k-means), il faut répondre à quelques questions : sont-ils des groupes ? Si oui, combien sont-ils ? C'est parce que tous les algorithmes de classification sera créés groupes quelle que soit le nombre de classes. Le processus de détermination du nombre de clusters est appelé tendance à l'évaluation visuelle VAT.

## II.2. Classification

### II.2.a Définition

La Classification consiste à regrouper les données en groupes homogènes appelés **classes** ou **clusters**, de sorte que les éléments à l'intérieur d'une même classe soient similaires, et les éléments appartenant à deux classes différentes soient différents [5]. Il faut donc définir une mesure de similarité entre deux éléments des données généralement on utilise la mesure distance. Choix d'une distance

Chaque élément peut être défini par les valeurs de ses attributs, ou d'un point de vue mathématique par un vecteur ou un point. Le nombre d'éléments de ce vecteur est le même pour tous les éléments il est appelé la dimension du vecteur. Etant donné deux vecteurs  $x_1$  et  $x_2$ , il faut définir la distance entre ces deux éléments  $d(x_1, x_2)$ . Parmi les choix les plus classiques, on retrouve :

Distance euclidienne :

C'est la distance entre 2 points. Etant donné deux points A et B repérés par leurs coordonnées X et Y, (Figure 14)

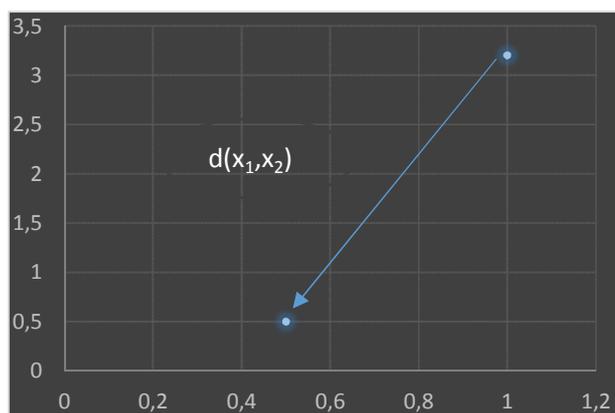


Figure 54 distances euclidiennes entre  $x_1, x_2$

on a :

$$d(A, B) = \sqrt{(X_B - X_A)^2 + (Y_B - Y_A)^2}$$

## Une Classe (cluster)

Un groupe est donc une collection d'objets qui sont « similaires » entre eux et sont « différents » pour les objets appartenant aux autres groupes

### II.2.d. Principe des algorithmes classification

Les algorithmes de classification consistent à assigner des pixels à des classes en respectant les règles suivantes (Figure 15) :

- La distance entre les éléments d'une même classe (distance intra-classe) est minimale.
- La distance entre chaque classe (distance interclasses) est maximale.

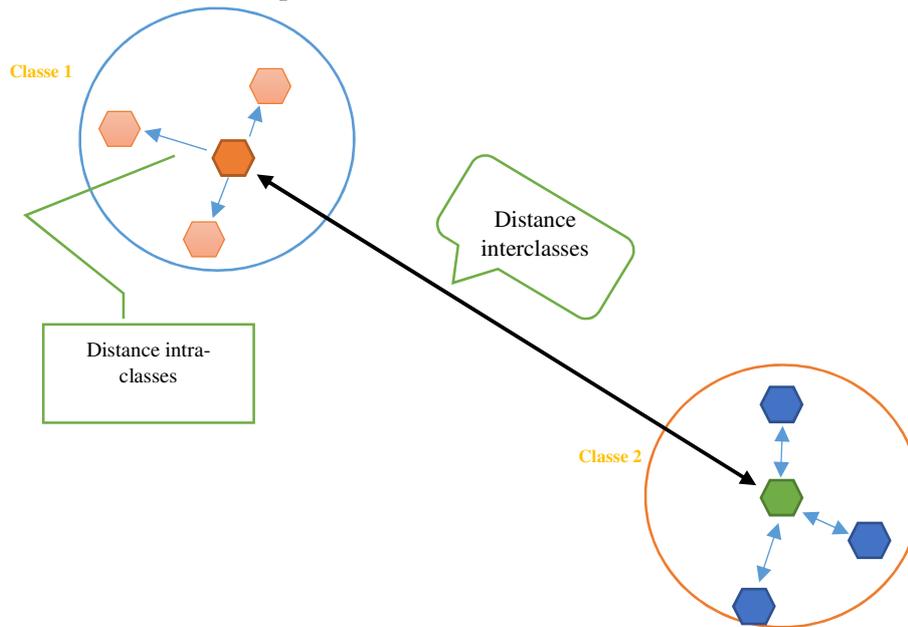


Figure 65 Principe de Classification

### II.2.c Méthodes de classification

#### Classification supervisée

##### Définition

Classification supervisée fournit une vue de groupée selon des catégories prédéfinies. C'est-à-dire l'assignation des éléments aux classes. Cela signifie que nous devons d'abord déterminer le nombre et la nature des classes que nous voulons utiliser [6].

Parmi les plus célèbres algorithmes dans classification supervisée on cite K-plus proches voisins.

##### Algorithme K-plus proches voisins

##### Définition

L'algorithme des  $k$ -plus proches voisins est un des algorithmes de classification les plus simples. Le seul outil dont on a besoin est une distance entre les éléments que l'on veut classifier. Si on représente ces éléments par des vecteurs de coordonnées, il y a en général pas mal de choix

possibles pour ces distances, partant de la simple distance usuelle (euclidienne) en allant jusqu'à des mesures plus sophistiquées pour tenir compte si nécessaire de paramètres non numériques comme la couleur, la nationalité, etc. Afin de trouver les  $K$  plus proches d'une donnée à classer on peut choisir la distance euclidienne.

### Notations [16]

Soit  $L = \{(x', c) | x' \in R^d, c \in C\}$  l'ensemble d'apprentissage

Soit  $x$  l'exemple dont on souhaite déterminer la classe

### Algorithme 1

#### Début

Pour chaque (exemple  $(x', c) \in L$ ) faire Calculer la distance  $D(x, x')$

#### Fin

Pour chaque  $\{x' \in kppv(x)\}$  faire compter le nombre d'occurrences de chaque classe

#### Fin

Attribuer à  $x$  la classe la plus fréquente;

#### Fin

### Avantages

- ✓ Simplicité, pas d'apprentissage d'un modèle
- ✓ Bonnes performances en général

### Des limites

- ✓ Paramétrage difficile (choix de la taille du voisinage)
- ✓ Nécessité de garder sous la main la base de données

## Classification non supervisée

### Définition

Classification non-supervisé consiste à tirer de la valeur de données dans lesquelles l'attribut à prédire n'apparaît pas [7].

### Méthode de partitionnement

Les algorithmes basés sur la représentation consistent à désigner un représentant pour chaque classe afin de calculer les distances plutôt que calculer des distances avec tous les éléments d'une même classe. Dans k-means, le représentant d'une classe est souvent le barycentre des points de cette classe (généralisation de la moyenne).

Le fonctionnement de l'algorithme est le suivant [9] :

- Initialisation : le nombre de classes  $K$  étant imposé, choisir  $K$  points aléatoirement pour constituer initialement les représentants de chaque classe
- Pour chaque point :
  - ✓ Calculer les distances entre ce point et les représentants des classes
  - ✓ Affecter à ce point la classe avec laquelle sa distance est minimale
  - ✓ Mettre à jour les représentants de chaque classe (par exemple, calcul de barycentre)

## La Classification Ascendante Hiérarchique

La classification ascendante hiérarchique (CAH) est une méthode de classification itérative dont le principe est simple [10].

1. On commence par calculer la dissemblance entre les  $N$  objets.
2. Puis on regroupe les deux objets dont le regroupement minimise un critère d'agrégation donné, créant ainsi une classe comprenant ces deux objets.
3. On calcule ensuite la dissemblance entre cette classe et les  $N-2$  autres objets en utilisant le critère d'agrégation. Puis on regroupe les deux objets ou classes d'objets dont le regroupement minimise le critère d'agrégation.

On continue ainsi jusqu'à ce que tous les objets soient regroupés.

Ces regroupements successifs produisent un arbre binaire de classification (dendrogramme), dont la racine correspond à la classe regroupant l'ensemble des individus. Ce dendrogramme représente une hiérarchie de partitions. On peut alors choisir une partition en tronquant l'arbre à un niveau donné, le niveau dépendant soit des contraintes de l'utilisateur (l'utilisateur sait combien de classes il veut obtenir), soit de critères plus objectifs.

Le fonctionnement de la classification hiérarchique est le suivant :

- Initialisation : chaque point est une classe
- Tant qu'il reste des points à classer :
  - ✓ *Calculer toutes les distances entre les points à classer et les représentants des classes*
  - ✓ *Fusionner les 2 classes les plus proches*
- Fusionner les classes les plus proches jusqu'à obtenir le nombre de classes voulu

Malgré les nombreux algorithmes qui ont été développés pour trouver un moyen ou un algorithme qui permet de classer des données automatiquement par la machine, cela s'est avéré très difficile et tous les algorithmes souffraient de lacune importante. Avant d'utiliser une méthode de classification on doit décider s'il existe des clusters, et si oui, combien sont-ils ?

### II.3. Tendance à évaluation visuelle VAT (visual assument tendency)

Les chercheurs ont depuis longtemps reconnu la difficulté de déterminer le nombre de groupe à rechercher dans un ensemble de données non étiquetées. En 1953, R. L. Thorndike [11] demandé, "Alors, comment allons-nous décider de la valeur de  $k$  - le nombre de familles ou de groupes? Et une fois  $k$  a été déterminé, comment allons-nous décider sur les limites et les cent roides des différents groupe?" Après avoir examiné ces questions, Thorndike [11] a poursuivi en disant "Commençons avec le deuxième problème Première [trouver les groupe], parce qu'il ressemble un peu plus docile et prête à l'attaque [Que de décider combien de rechercher] ..." [11] La technique que Thorndike employées (et qui est encore fréquemment utilisé aujourd'hui) était d'énumérer candidat solutions par application d'un algorithme de regroupement de données pour différentes valeurs de  $k$ . Se référant à cette technique, Thorndike a déclaré, [11]"on voudrait quelque type de test de signification du changement dans le critère lorsque  $k$  augmente de 2 à 3 et ainsi de suite. En effet, Thorndike a soulevé deux questions liées à la mise en groupe Comment peut-on évaluer le nombre de groupe dans un ensemble de données avant la mise en groupe ?

Comment peut-on valider des solutions produites par un algorithme de regroupement?

L'auteur Bezdek [12] a développé l'algorithme VAT pour la classification des groupes pour répondre à des questions précédentes. Mais Dans certains cas cette algorithme ne fonction pas mieux. Il a également développé une autre version iVAT (VAT amélioré) pour les cas des données du chemin d'accès de base. Par conséquent, nous pouvons utiliser l'une de l'algorithme de visualisation entre les trois selon notre degré de la complexité des données.

#### II.3.a. Définition

L'algorithme de VAT est une méthode visuelle pour déterminer le nombre possible de clusters dans un ensemble d'objets.

**Accès visuel tendance** Accès visuel tendance La procédure suivie dans l'algorithme VAT est donnée dans la Figure suivant :

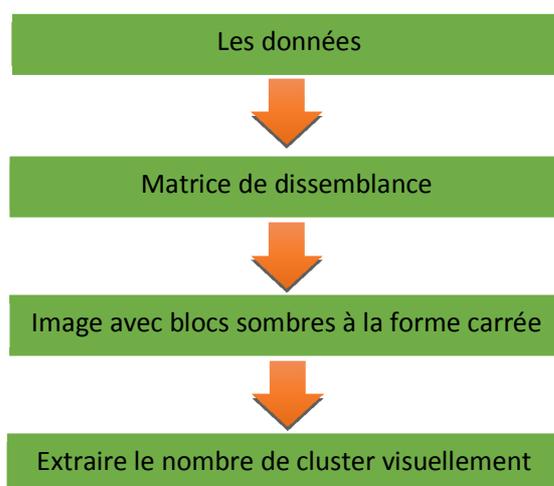


Figure 16 Accès Visuel Tendence

### II.3.b. Matrice de dissemblance et image de VAT

La matrice de dissemblance  $R$  est l'entrée de VAT, elle est transformée en entrée en image ( $I(R)$ ) composée de blocs sombres en forme de carré (Figure 17). Ainsi, l'algorithme VAT génère en sortie l'image ( $I(R')$ ) sur laquelle nous pouvons détecter le nombre de groupes par comptage de blocs sombres en forme de carrés le long de la diagonale de l'image (Figure 18).

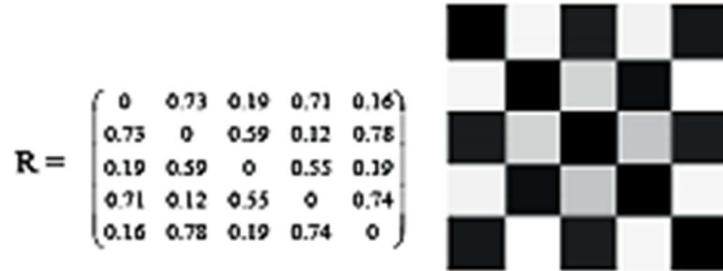


Figure 17  $I(R)$

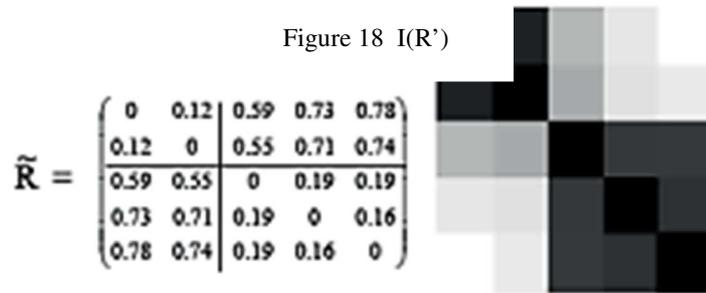


Figure 18  $I(R')$

#### Notation

- Objets  $O = \{o_1, \dots, o_N\} : o_i = \text{objet physique}$
- Les données d'objet  $X = \{x_1, \dots, x_N\} : x_i = \text{vecteur de caractéristiques pour } o_i$
- $x_{ji} = j$  le (mesure) caractère of  $x_i : 1 \leq j \leq p$
- Les données relationnelles  $R = [R_{ij}] = \text{object}(o_i, o_j) \text{ or } (x_i, x_j)$
- $S_{ij} = \text{similitude}(o_i, o_j) \text{ or } (x_i, x_j)$
- $D_{ij} = \text{dissimilitude}(o_i, o_j) \text{ or } (x_i, x_j)$
- Nous supposons (i)  $R_{ij} \geq 0 \quad 1 \leq i \neq j \leq n$
- ( $R = d$ ) (ii)  $R_{ij} = 0 \quad 1 \leq i \leq n$
- (iii)  $R_{ij} = R_{ji} \quad 1 \leq i \neq j \leq n$

Lorsque  $X \rightarrow D, R_{ji}=D_{ij} = \|x_i - x_j\|$  est euclidienne

### II.3.c. Algorithme De VAT

**Entrée :**  $n \times n$  matrice des dissemblances  $R = [r_{ij}]$  satisfaisant,

Pour tout  $1 \leq i, j \leq n : r_{ij} = r_{ji}, r_{ij} \geq 0, r_{ii} = 0$ .

**Étape 1.** Mettre  $I = \emptyset; J = \{1, 2, \dots, n\}; P = (0, 0, \dots, 0)$

Sélectionner  $(i, j) \in \operatorname{argmax}\{r_{pq}\}$

Mettre  $P(1) = j$

Remplacer  $I \leftarrow I \cup \{j\}$  et  $J \leftarrow J - \{j\}, p \in J, q \in J$

**L'étape 2.** Pour  $t = 2, \dots, n :$

Sélectionnez  $(i, j) \in \operatorname{argmin}\{r_{pq}\}$

Mettre  $P(t) = j$

Remplacer  $J \leftarrow I \cup \{j\}$  et  $J \leftarrow J - \{j\}, p \in J, q \in J$

Suivant  $t$ .

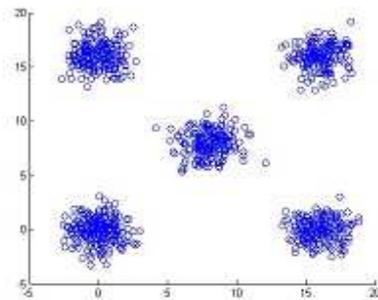
### Étape 3.

Former la matrice de dissimilitude commandée  $R^* = [r_{ij}^*] = [r_{p(i)p(j)}]$  pour  $1 \leq i, j \leq n$ .

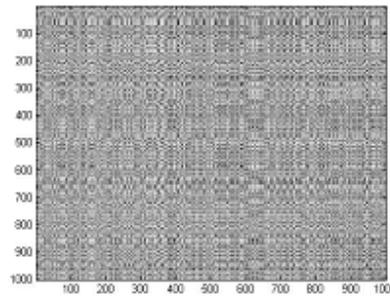
**Sortie:** Image  $I(R^*)$ . Échelle de sorte que  $\max_{i,j}\{[r_{ij}^*]\}$  Correspond à blanc et  $\min_{i,j}\{[r_{ij}^*]\}$  En noir.

### Exemple

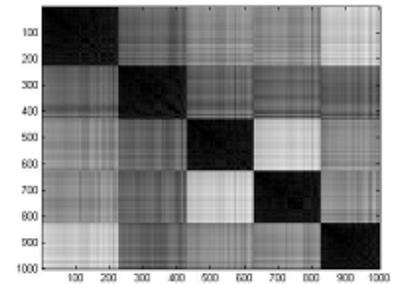
La Figure (19) est un diagramme de dispersion de  $n = 1000$  .Tirées d'un mélange de cinq distributions normales. Les moyens, proportions de mélange. Et le nombre d'échantillons dans chaque groupe (c'est-à-dire, la cardinalité  $n_i, i = 1, 2, 3, 4, 5$ ).



(a) Diagramme de dispersion de données d'entrée



(b)  $I(R)$



(c)  $I(R^*)$

Figure 18 : Un exemple de la VAT qui indique  $c = 5$

### **II.3.d. Avantage et les limite de l'outil le VAT**

#### **Avantages**

- Il présente les informations en mode Visuel, même un utilisateur normal peut calcule le nombre de clusters.
- Est utilisée cette visualisation à n'importe quel type de données telles que le texte, image, audio et vidéo etc. pour définir le nombre des groupes ou des classes

#### **Limite**

- Il prend beaucoup de temps pour résoudre des  $k$  valeurs où dans le cas du données volumineuses, Dans de nombreux cas ne donnent pas des solutions
- les résultats visuels ne sont pas clairs si les clusters se chevauchent.

Dans certains cas, le VAT ne fonctionnent pas biens, parce que les images peuvent se remplis De nombreux de l'information donc VAT produit des photos ( $I(R')$  vague et fournis des carrées sombre non détectable, cela peut résolu par l'utilisant VAT amélioré (IVAT)

## II.4. Version Amélioré VAT (improved VAT (iVAT))

### II.4.a Définition

La iVAT remplace chaque dissemblance  $r_{ij}$  dans  $R$  Par le chemin minimax de distance entre les nœuds  $i$  et  $j$ . (Objets  $i$  et  $j$  dans les données d'entrée)[13].

Laissez  $P_{ij}$  l'ensemble de tous les chemins possibles de  $o_i$  à  $o_j$ . Le chemin en fonction distance entre  $o_i$  et  $o_j$  Est considéré dans la transformation :

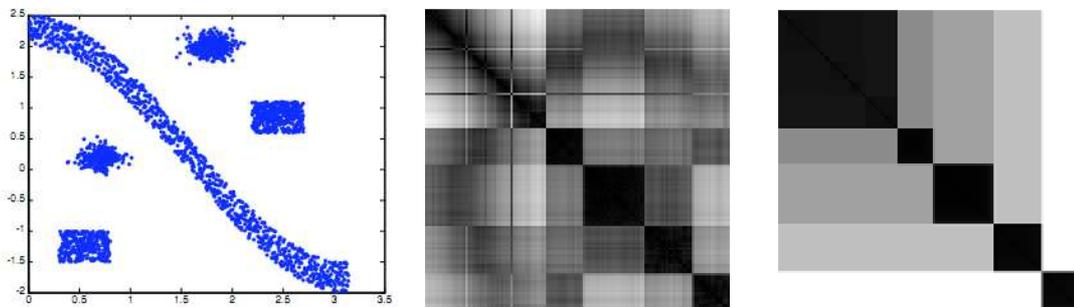
$$R' = [r'_{ij} = \min_{p \in P_{ij}} \{ \max_{1 \leq h < |p|} \{ r_{p[h]p[h+1]} \} \}; \{ i, j \leq n \}] \quad (1)$$

Où  $p[h]$  indique l'objet à la position  $h$  dans le chemin  $p$ . Et  $|p|$  est la longueur de ce chemin.

L'exemple suivant (Figure 20) affiche un diagramme de dispersion des données contenant 5 classes après application de l'algorithme VAT. L'image résultante (Figure 19.b) n'a pas de structure visuelle claire. À l'inverse de l'image résultante de l'exemple précédent (Figure 19).

Le problème dans l'ensemble de données de Figure 19, est la grande forme irrégulière située dans le Centre des données, qui occupe la partie supérieure gauche de l'image la VAT.

L'image  $C$  dans la Figure 19 est l'image d'iVAT  $I(R'^*)$  de cet ensemble de données, faite par la première transformation  $R \rightarrow R'$  Avec le chemin en fonction transformée de distance en (1), et puis en appliquant le VAT à  $R'$ . L'image iVAT montre clairement la structure sous-jacente des données d'entrée.



(a) Groupes irréguliers dans deux dimensions

(b)  $I(R^*)$

$I(R'^*)$

Figure 20 VAT et iVAT

### II.4.b Algorithme de iVAT

**Entrée :**  $R^*$  - (VAT - réordonnée dissimilarités matrix)

**Initialiser :**  $R'^* = [0]^{n \times n}$

**Pour**  $r = 2, \dots, n$  **faire**

$$j = \operatorname{argmin}\{R_{rK}^*\}; K = 1, \dots, r - 1$$

$$R'_{rc} = R_{rc}^*, c = j$$

$$R'_{rc} = \max\{R_{rj}^*, R'_{jc}\}, c = 1, \dots, r - 1 \dots c \neq j$$

**Pour**  $2 \leq j \leq n; i < j$  faire

$$R'_{ij} = R'_{ji}$$

$$R'^* = (R'^*)^T \rightarrow [R'^*]_{rc} = [R'^*]_{cr}$$

**Sortie:**  $D'^*$

Nous supposons que

$$\max_{1 \leq i, j \leq n} \{d_{ij}^*\} = 0 = \text{noire}$$

$$\max_{1 \leq i, j \leq n} \{d'_{ij}\} = \text{Blanc}$$

-Nous allons étudier 2 méthode qui visant à trouver automatiquement le nombre de clusters : CCE et DBE.

## II.5. Extraction de nombre de cluster (Cluster Count Extract (CCE))

### II.5.a Définition

CCE (Cluster Count Extraction) est un algorithme d'extraction automatique de nombre de clusters. Une idée immédiate pour considérer un algorithme de détection de bord est de trouver les bords verticales (fig.19.c) (ou horizontales) des blocs sombres [14].

### IV.5.b Algorithme

Les notions suivant sont utiles pour la coopération de l'algorithme CCE :

**Otsu** : la méthode d'Otsu est utilisée pour effectuer un seuillage automatique à partir de la forme de l'histogramme de l'image, ou la réduction d'une image à niveaux de gris en une image binaire [15].

**FFT** :

Méthodes FFT est La Transformée de Fourier Rapide est un algorithme qui convertit une fonction du temps à valeurs complexes échantillonnées en une fonction à valeurs complexes de la fréquence.

**Entrée** :  $N \times N$  l'Image  $I(R')$ , on suppose pour que  $\max = \text{blanc}$  et  $\min = \text{noir}$

**Etape 1** : Seuil  $I(R')$  avec la méthode d'Otsu

**Etape 2** : Produisez le rapport de filtre de corrélation de grandeur  $s$ .

**Etape 3** : Appliquer la FFT à l'image de VAT segmenté et le filtre.

**Etape 4** : Multiplier l'image transformée de VAT avec le conjugué complexe du filtre transformé.

**Etape 5** : Calculer l'inverse FFT de l'image filtrée.

**Etape 6** : Calculer histogramme des valeurs (pixel) hors diagonale de l'arrière image transformée.

**Etape 7** : Coupez l'histogramme à une ligne horizontale arbitraire  $y = b$  (généralement  $b = 0$ ) et comptez le nombre de pointes.

**Sortie** : entier (C), une estimation du nombre de blocs sombres le long de la diagonale du  $I(R')$   
CCE a le potentiel de remplacer la nécessité d'une interprétation humaine du nombre de clusters dans un ensemble de données [14].

## IV.6. Méthode L'extraction de bloc foncé (Dark Block Extraction(DBE))

### IV.6.a Définition

DBE est une méthode d'estimation automatique de nombre de blocs sombres (probablement aussi le nombre de groupes potentiels). Combinant plusieurs techniques de traitement du signal. Le DBE accepte un  $N \times N$  ordonné et mis à l'échelle de dissemblance  $R = [r_{ij}]$  et un paramètre  $a$ . la proportion de la taille de cluster minimale autorisée pour la taille des données  $N$ , comme entrées. Étapes de l'algorithme de la DBE sont données ci-dessous.

### IV.6.b Algorithme DBE

**Étape 1 :** Transformer  $R$  à une nouvelle matrice de dissemblance  $R'$  Avec  $r'_{ij} = 1 - \exp(-\frac{r_{ij}}{\sigma})$

Où  $\sigma$  est un paramètre d'échelle déterminée automatiquement à partir de  $R$  en utilisant l'algorithme de Otsu.

**Étape 2 :** Former une image  $RDI$  à l'aide d'algorithme de VAT.

**Étape 3 :** Seuil  $I(R^*)$  pour obtenir l'image binaire  $I(R_2^*)$  à l'aide de l'algorithme d'Otsu.

**Étape 4 :** Filtrer  $I(R_2^*)$  à l'aide d'opérateurs morphologiques avec ligne directionnelle élément structurel de taille  $l_1 = aN$ , Pour obtenir l'image binaire filtrée de  $I(R_3^*)$ .

**Étape 5 :** Effectuer transformée de distance sur  $I(R_3^*)$  Pour obtenir de la nouvelle image en échelle gris de  $I(R_4^*)$ . L'échelle des valeurs de pixels à  $[1.0]$ .

**Étape 6 :** Projet des valeurs de pixel de  $I(R_4^*)$  sur l'axe principal de diagonale de l'image à la projection de la forme du signal  $H_1$ .

**Étape 7 :** Lisser le signal  $H_1$  pour obtenir le signal filtré  $H_2$ , utilisation le filtre moyenne de longueur  $l_2 = a2N$

**Étape 8 :** Calculez le premier dérivé d'ordre de  $H_2$  to obtiennent  $H_3$ .

**Étape 9 :** Trouvez le  $p_i$  de position maximal (conforme aux points traversant le zéro positifs-à-négatif)

Et position de la vallée  $v_i$  (correspondant au négatif à positif points de passage par zéro) dans  $H_3$ .

**Étape 10 :** Compter les pics majeurs en retirant le mineur à l'aide d'un filtre de taille  $l_3 = a2N$

## II.7. Conclusion

L'algorithme VAT vise à trouver le nombre de classes dans un nuage de données, pour cette algorithme il y a plusieurs variantes on peut citer le iVAT.

L'application de l'algorithme VAT ou iVAT donne en sortie une image comportant un nombre de blocs sombres qui représentent le nombre de clusters.

Afin d'extraire le nombre automatiquement et éviter l'intervention de l'utilisateur les algorithmes CCE ou DBE peuvent être appliqués.

# **CHAPITRE 3**

**Conception ET implementation**

### **III.1. Introduction**

Après avoir pris connaissance dans les chapitres 1 et 2 la méthode Tendance à évaluation visuelle (VAT) ainsi que ses variantes, ce chapitre porte sur la construction de l'application.

Le but de l'application est de calculer le nombre de centre des images en général et en particulier les images satellitaires.

Dans un premier temps nous allons présenter les ressources utilisées, Après cela, nous allons discuter les résultats obtenus

### **III.2. Ressources Utilisées**

#### **III.2.a. Les ressources matérielles :**

- Processeur I5 4440 3.3 GHZ
- Une mémoire vive d'une capacité de 4Go .
- Une carte graphique de 1Go.

#### **III.2.b. Les ressources logicielles :**

- Système d'exploitation : Windows 7.
- Langage de programmation MATLAB.
- L'éditeur utilisé est le MATLAB R2010a .

Matlab est un logiciel pour effectuer des calculs numériques. Il a été conçu initialement pour faciliter le traitement des matrices mais il est maintenant utilisé dans tous les domaines des sciences qui nécessitent de faire des calculs.

Le langage contient de nombreux avantages par rapport à autre langages [20]:

- Une librairie très riche
- La continuité parmi les valeurs entières, réelles et complexes,
- La programmation facile, et très efficace, optimisé pour le traitement des matrices,
- Programmation infiniment plus rapide pour le calcul et pour l'affichage
- Une aide très bien faite
- Code facile à comprendre et très lisible
- Possibilité d'exécuter du code en dehors du programme
- Possibilité d'inclure un programme en C/C++

### **III.3. Conception**

Le but principal de notre application est de calculer le nombre de classe ou de centre pour n'importe quelle image, pour terminer le processus nous passons plusieurs étapes:

- ✓ Chargement de l'image.

- ✓ Calculer VAT et la présentée sous la forme d'image avec algorithme VAT ou iVAT et la présentée sous la forme d'image avec algorithme iVAT
- ✓ Extraire le nombre de centres de l'image précédente sous la forme de nombres entiers avec CCE
- ✓ Affichage du résultat obtenu

Ces étapes peuvent se résumer par l'organigramme suivant :

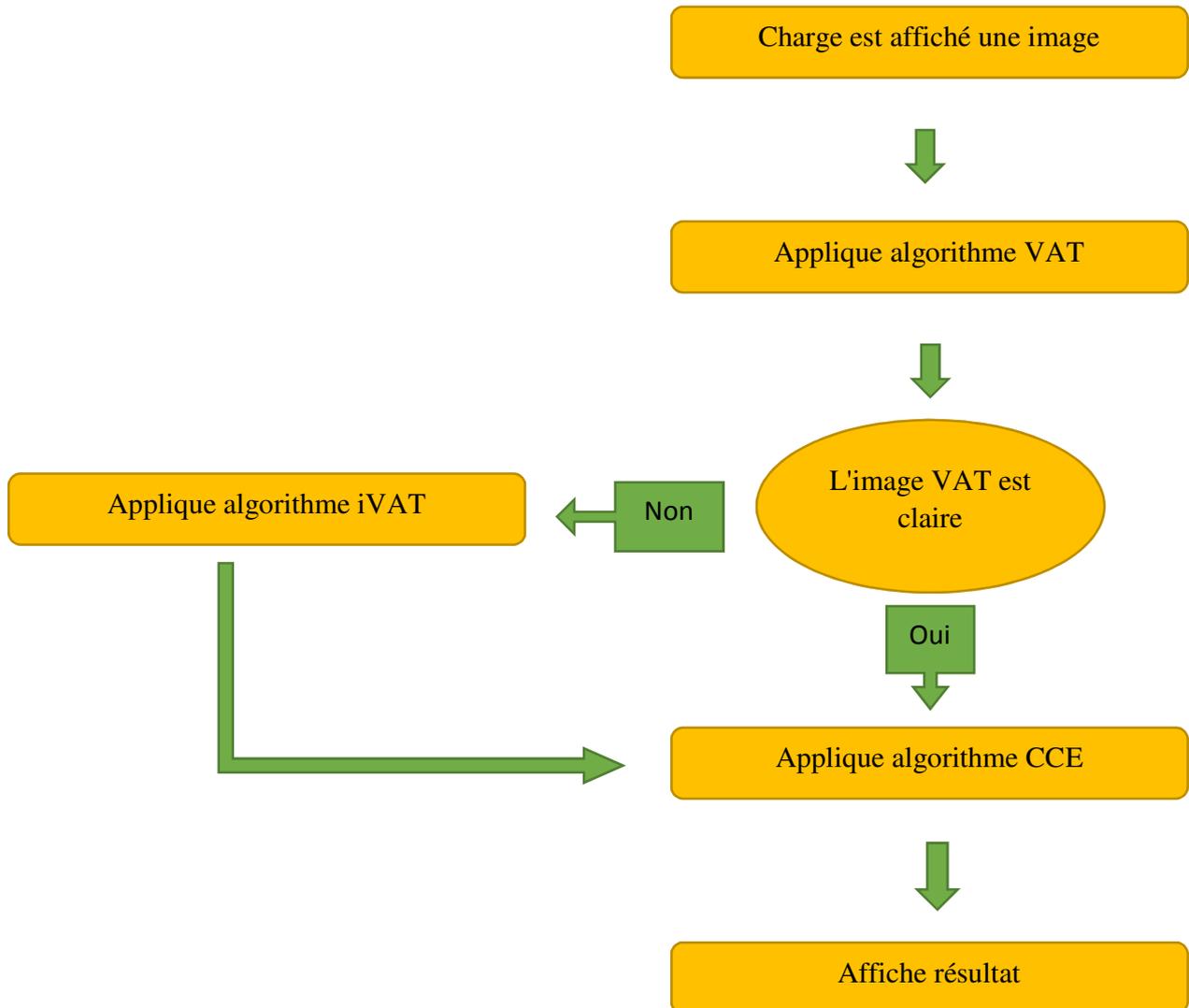


Figure 21 Organigramme de l'application.

### III.4. Données utilisées

Pour tester notre application différentes images ont été utilisées, on peut les divisées en deux catégories :

**Les images nature** : Nous avons utilisé une image connue dans la communauté du traitement d'image à savoir Peppers (Figure 22).

**Les images satellitaires** : Nous avons opté pour l'utilisation des images satellitaire

- ✓ **Mosquée Hassan II, Casablanca, Maroc** et elle a été acquise par le satellite Pléiades 1A 2011.( Figure 23 )[17].
- ✓ l'irrigation vue par le satellite Pléiades à Olton 22 février 2012(Figure 24) [18].
- ✓ des Iles Fidji par satellite SPOT-7 14 août 2014 (Figure25) [19].

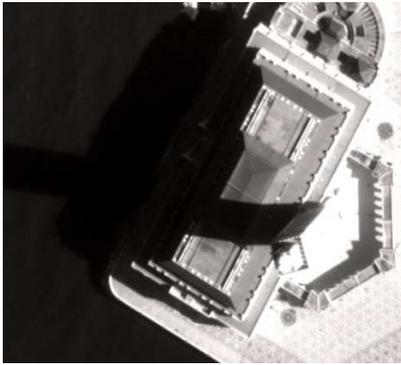


Figure23 Mosquée Hassan II, Casablanca,



Figure22 Image 'peppers'

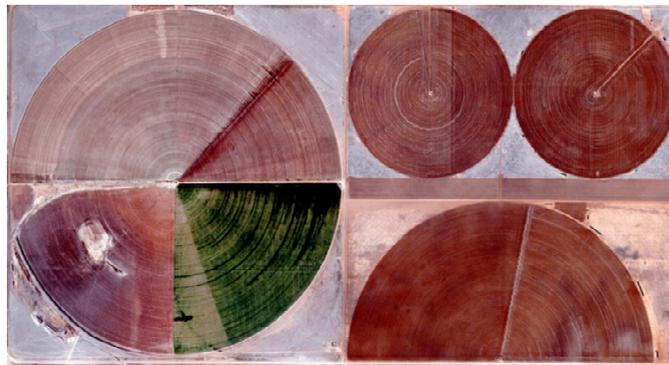


Figure 24L'irrigation à Olton vue par le satellite Pléiades



Figure 25 Les Iles Fidji par SPOT-7

### III.5. Présentation de l'application

On commence la présentation de notre application par l'interface d'accueil.

Notre application permet de :

- Charger les images à partir de ses emplacements et de les afficher.
- Appliquer l'algorithme VAT
- Appliquer l'algorithme iVAT
- Appliquer l'algorithme CCE

L'interface principale de l'application est représentée par la figure suivante (Figure 26) :



Figure 26 L'interface principale de

On charge les images (Ctrl+I)



Figure 27 Chargement des

La matrice similarité

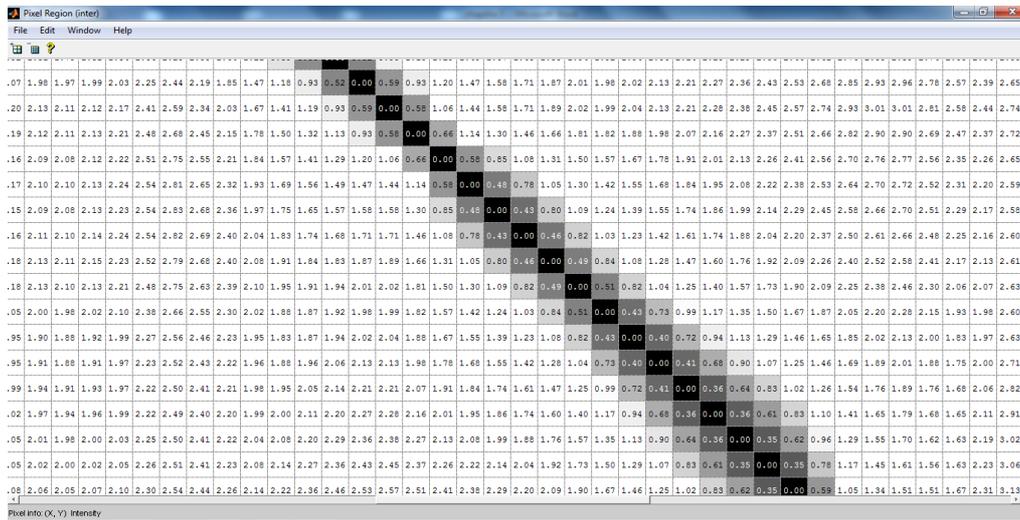


Figure 28 matrice de similarité

### III.5.a. Application d'algorithme VAT

Après le chargement de l'image, on clique sur Botton VAT afin d'appliquer l'algorithme VAT, l'image suivante sera affichée



Figure 29 L'interface d'algorithme VAT

Après cela on applique l'algorithme CCE en appuyant sur la « VAT\_cluster », le nombre de classes sera affichée (Figure 30).

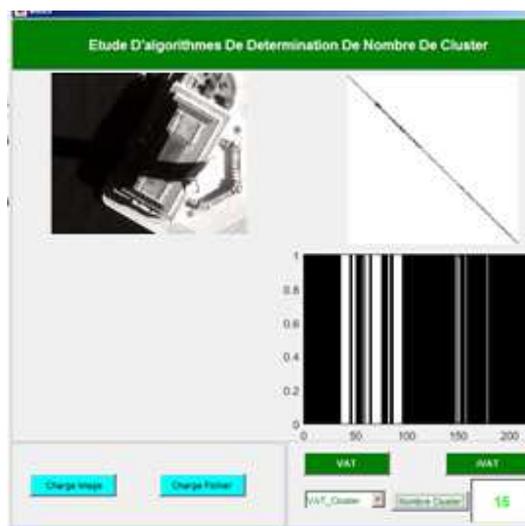


Figure 30 Application l'algorithme CCE

### III.5.b. Application d'algorithme iVAT

Pour faire le déroulement l'iVAT en charge une image comme mentionné précédemment, On clique sur VAT et ensuite sur le bouton iVAT :

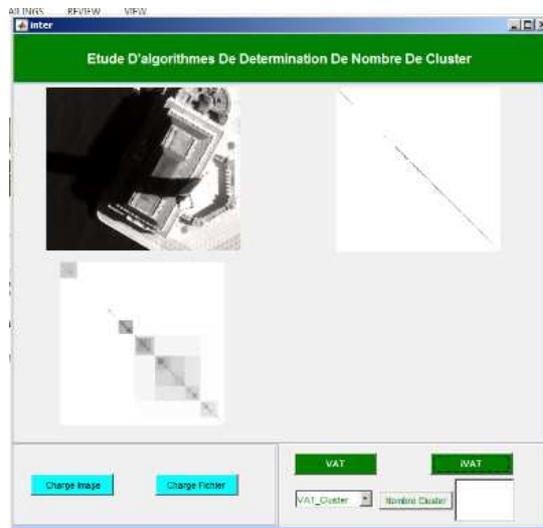


Figure 31 Application l'algorithme iVAT

Image Otsu de iVAT

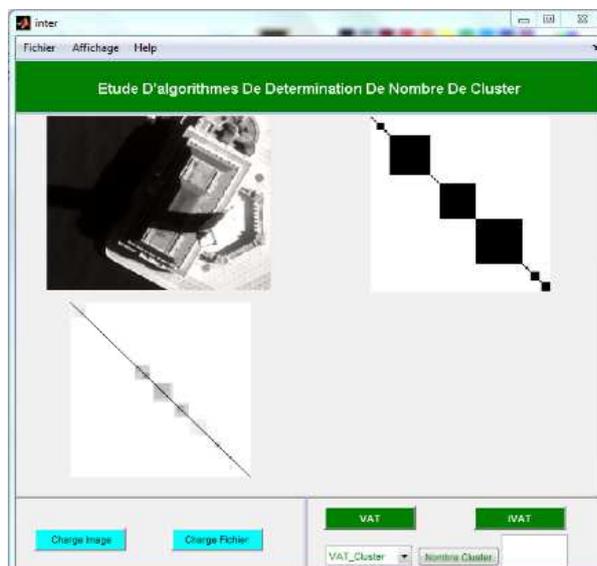


Figure 32 Otsu de iVAT

Pour affiche le nombre de center on clique sur « iVAT\_cluster »

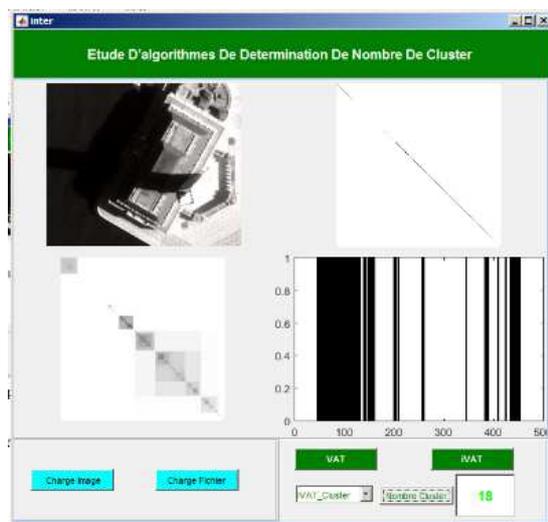


Figure 33 Interface d'iVAT

Aussi, le résultat image de VAT ou iVAT peut être enregistré :

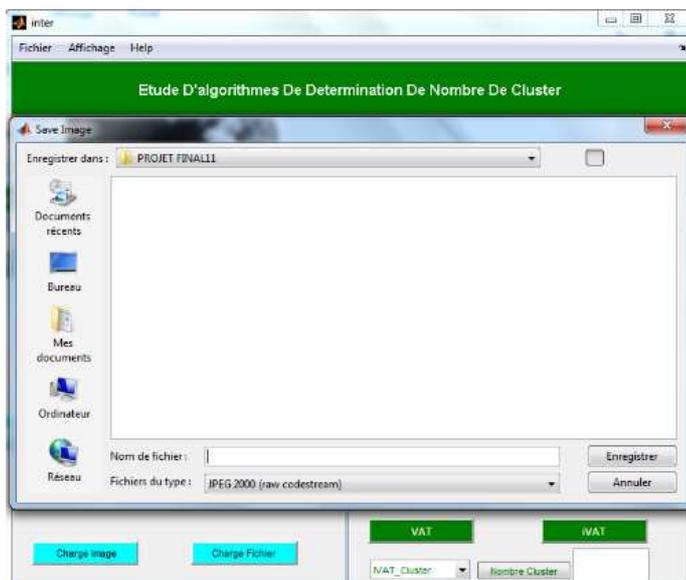
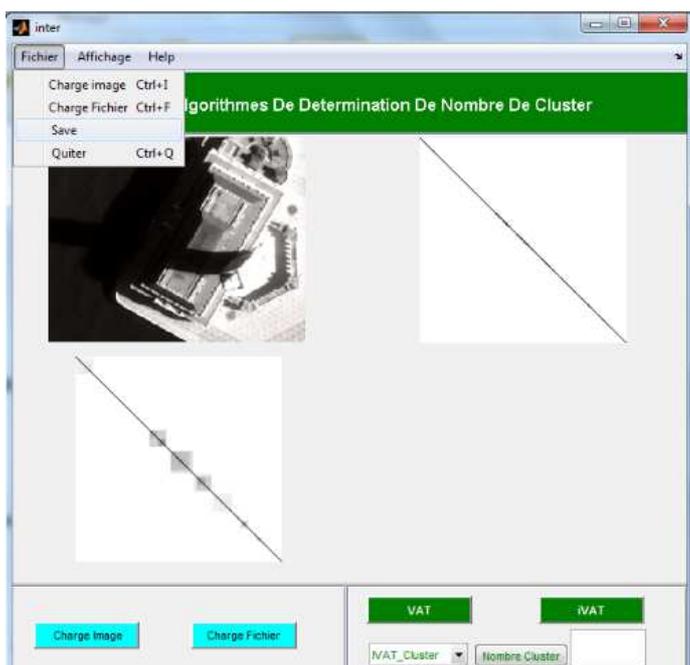


Figure 34 sauvegardes résultat

Image	Nombre de centres VAT	Nombre de centres iVAT	Temps de calcul (s) VAT	Temps de calcul (s) iVAT
Figure 22	15	20	0.5	1.8
Figure 23	19	18	4.2	5
Figure 24	10	37	6	7.7
Figure 25	8	14	4	3,6

Figure 35 Le nombre de groupes en termes de temps

Pour tous les jeux de test effectués, on remarque toujours que l'image résultante n'est pas claire, la différence entre l'algorithme VAT et iVAT réside dans le nombre de centres ou le temps de calcul.

### III.6. Discussions

Les résultats obtenus montrent que les images obtenues à partir de la méthode VAT sont sensibles aux détails de l'image à traiter : Si les couleurs se chevauchent entre elles le résultat sera toujours la valeur 0, mais si les couleurs sont bien séparées on aura des résultats raisonnables.

L'algorithme iVAT donne relativement de bon résultats que son prédécesseur, mais quand les couleurs se chevauchent, ils ne donnent aucun résultat comme l'algorithme VAT.

On remarque aussi que si la résolution de l'image augmente l'efficacité diminue et cela pour les deux algorithmes.

### **III.7. Conclusion**

Dans ce chapitre nous avons présenté l'implémentation et la conception de notre logiciel, en jouant sur les différents algorithmes et les différentes images.

L'algorithme de VAT donne de bons résultats dans certaines circonstances avec une exécution rapide. L'algorithme iVAT donne un meilleur résultat mais prend beaucoup plus de temps à exécuter. Dans le cas images à grandes dimensions les deux algorithmes ont échoué à donner le nombre de centres adéquat, cela peut être résolu par l'utilisation d'autres variantes comme bigVAT et soVAT.

## **Conclusion générale**

Dans ce mémoire, Nous avons étudié le problème de l'extraction de nombre de centres dans les images satellitaires. Ce problème est devenu un obstacle majeur surtout lorsqu'on étudie des zones inconnues, tels que les zones d'extraction. Les algorithmes VAT et iVAT ont été implémenté dans le cadre de ce projet afin de résoudre ce problème.

Nous somme passé par trois chapitres pour réaliser ce projet :

Chapitre I : Nous avons donné une définition complète et détaillée sur la télédétection et l'imagerie.

Chapitre II : On a présenté une introduction sur la classification avec mention des types les plus importants et les problèmes rencontrés. Nous avons aussi survoler les méthodes VAT, iVAT et CCE.

Chapitre III : On a présenté l'application ainsi que les différents tests effectués et les résultats trouvés

Comme perspectives nous proposons d'étaler ce projet vers l'implémentation des autres variantes du VAT (BigVat et soVat) afin d'élargir le type et la dimension d'images utilisées dans notre application.

## Bibliographie

- [1] Bonn F. et Rochon G, livre , « Précis de télédétection ».1992.
- [2] Baret F., 1986 ,livre,« Contribution au suivi radiométrique de cultures de céréales ». Université de Paris.
- [3] Baret F., Guyot G. and Major D ,Thèse, « La biomasse de cultures à l'aide d'évaluation des mesures radiométriques. Photogrammétrie »,1990.
- [7] E. Lebarbier, T. Mary-Huard . AgroParisTech. These,« Classification non supervisée »,1989.
- [8] JANNADI Wiem-TURKI Imen. These,« Implementation of K-Means Clustering in Cloud Computing Environment Réalisé »,2001,Paris.
- [11] Thorndike, R.L. livre.« Who belongs in the family? » 267–276 (1953).
- [12] J.C. Bezdek- R.J. Hathaway , article.« VAT: A Tool for Visual Assessment of (Cluster) Tendency»,2007.
- [13] Liang Wang-Uyen T. V. Nguyen-James C. Bezdek-Christopher A. Leckie-and Kotagiri Ramamohanarao«iVAT and aVAT: Enhanced Visual Analysis for Cluster Tendency Assessment». Département de Science Informatique et de Génie Logiciel de L'Université de Melbourne, Victoria, 3010, Australie-
- [14] Isaac J. Sledge-Jacalyn M. Huband -James C. Bezdek « (Automatic) Cluster Count Extraction from Unlabeled Data Sets»-Cinquième Conférence Internationale sur les Systèmes Flous et de Découverte de Connaissances
- [16] jérome azé,These, K-plus proches voisins- - Université Paris 11 / LRI-CNRS- 1er mars 2007

## Les Références électronique

- [4] <http://www.asal.dz/>
- [5] Dr. M.H. Dunham - <http://engr.sm....rt2.ppt>.
- [6] <http://grass.osgeo.org/.../book1.html>
- [9] <http://www.hypertextbo...../page001.html>
- [10] <http://rstat.ouvato.....rarchique>
- [15] [http://fr.wikipedia.org/wiki/M%C3%A9thode\\_d%27Otsu](http://fr.wikipedia.org/wiki/M%C3%A9thode_d%27Otsu)
- [17] <http://www.geo-airbusds.com/fr/2927-premieres-images-du-satellite-pleiades-1a>
- [18] <http://image-cnes.fr/1-loeil-du-satellite/agriculture-au-texas-l-irrigation-vue-par-le-satellite-pleiades-a-olton/>
- [19] <http://image-cnes.fr/1-loeil-du-satellite/spot-7-1eres-%C2%AB-cartes-postales-%C2%BB-de-la-terre/>
- [20] <http://www.samuelboudet.com/fr/matlab>

