

Fiche TP3 : Traitement

Objectif

Le traitement d'image satellitaire recouvre l'ensemble des techniques consistant à modifier une image dont le but de faciliter son interprétation ou d'extraire des informations d'ordre quantitatif.

L'objectif de TP3 est implémenté quelques méthodes de traitement d'image satellitaire telles que le rehaussement de contours, la segmentation et la classification automatique (en anglais clustering).

Pour l'application nous allons utiliser une partie de l'images satellitaire pour les deux composition vraies couleurs et fausses couleurs. Cette image est acquise par le satellite LANDSAT 7 ETM+ (Enhanced Thematic Mapper plus) des zones côtières de Sumatra, en Indonésie, la date d'acquisition est 14 Mai 2002.

Programme

Les différents traitements appliqués sur les images

Etape 1 : Rehaussement de contours

Le but de la **détection de contours** est de repérer les points d'une image numérique qui correspondent à un changement brutal de l'intensité lumineuse. Ces changements de propriétés de l'image traduisent en général des événements importants ou des formes qu'il est important de détecter. Certaines méthodes de détection de contours utilisent des filtres (linéaires ou pas), qui vont permettre de mettre en évidence les contours : on appelle cette tâche le **rehaussement de contours**.

Nous allons étudier ici le problème du rehaussement des contours grâce au filtrage linéaire. Ces filtres sont définis par leur masque $h(m, n)$ de taille $M_h \times N_h$. Ces filtres vont être conçus pour mettre en évidence des changements brutaux d'amplitude dans une direction donnée. Il y a deux types de filtres :

- Les filtres du premier ordre (**Prewitt** et **Sobel**) sont basés sur un calcul de dérivée. Les contours (dans la direction définie par le masque) sont visibles dans l'image filtrée car ils correspondent aux pixels ayant des amplitudes élevées (en valeur absolue)
- Les filtres du second ordre (**Laplacien** et **LOG**) sont basés sur un calcul de dérivée seconde. Cette fois-ci, les contours correspondent dans l'image filtrée à des transitions entre des valeurs négatives et positives.

1- **Créer et appliquer les filtres sur l'image satellitaire** (les trois niveaux de gris RVB pour chaque composition).

- 2- Pour les filtres du premier ordre, créer les fonctions permettant de calculer l'amplitude du gradient et la direction du gradient.
- 3- Observer et comparer l'effet de ces filtres sur différentes images (RVB).

Etape 2 : Segmentation

La **segmentation d'image** est une opération de traitement d'images qui a pour but de rassembler des pixels entre eux suivant des critères prédéfinis. Les pixels sont ainsi regroupés en régions, qui constituent un pavage ou une partition de l'image. Il peut s'agir par exemple de séparer les objets du fond.

Il existe de nombreux types de méthodes de segmentation d'image : certaines sont basées sur les contours, d'autres sur un seuillage des pixels en fonction de leur intensité, d'autres découpent l'image en régions connexes, etc....

Segmentation sur seuillage :

Le seuillage a pour objectif de segmenter une image en plusieurs classes en n'utilisant que l'histogramme. On suppose donc que l'information associée à l'image permet à elle seule la segmentation, i.e. qu'une classe est caractérisée par sa distribution de niveaux de gris. A chaque pic de l'histogramme est associée une classe. Il existe de très nombreuses méthodes de seuillage d'un histogramme. La plupart de ces méthodes s'appliquent correctement si l'histogramme contient réellement des pics séparés.

- 1- Créer une approche choisie parmi les 2 méthodes proposées pour la segmentation des deux images satellitaires RVB.
- 2- Vous préciserez tous les paramètres que vous utiliserez en expliquant leur sens (comment vous avez trouvé les seuils qui séparent au mieux les objets, quel histogramme vous avez choisi, ...) et l'effet qu'ils ont sur le résultat de la segmentation.
- 3- Quelles sont les limites de cette approche ?

Etape 3 : Classification

La **classification** consiste à classer chaque pixel dans une classe ou une autre. Chaque classe représente une région homogène. Cette approche est souvent utilisée pour isoler un objet du fond ou parfois plusieurs objets d'une scène. Deux variantes de cette approche existent : l'une est dite supervisée et l'autre dite non supervisée (ou automatique).

Exemple Introductif a la classification automatique d'une image couleur

Considérons par exemple une image couleur, chaque image contient n pixels (x_1, \dots, x_n) , chaque pixel x_i contient $d = 3$ valeurs (RGB). On peut donc représenter donc le i ème pixel ($i = 1, \dots, n$) par un vecteur x_i de dimension $d = 3$: $x_i = (x_{i1}, x_{i2}, x_{i3})^T \in \{0, 1, \dots, 255\}^3$. Si l'on connaît les classes de certains pixels, on pourra prédire les classes des autres pixels, on choisissant une mesure de (dis)similarité, par exemple une simple distance, ou une mesure de probabilité, etc. Chaque pixel à classer aura donc la classe de celui qui lui est le plus proche au sens de la mesure de (dis)similarité choisie. Ceci peut être utilisé par exemple en segmentation d'image. De manière

générale, on peut représenter les données comme un ensemble de vecteurs (x_1, \dots, x_n) , chaque x_i est composée de d composantes réelles : $x_i = (x_{i1}, x_{i2}, x_{i3})^T \in \mathbb{R}^d$.

- 1- Implémenter l'algorithme des centres mobiles (K-means ou K-moyenieur, ISODATA, Kmédiane, K-médoïd) pour la classification automatique d'un ensemble de données (x_1, \dots, x_n) des deux images satellitaires RVB. (Mettez la référence de l'approche).
- 2- Vous préciserez tous les paramètres que vous utiliserez en expliquant leur sens et l'effet qu'ils ont sur le résultat de la classification.
- 3- Quelles sont les limites de cette approche ?

Annexes

Filtre Prewitt

Mathématiquement, le filtre est composé de deux matrices 3×3 que l'on va convoluer avec l'image originale pour calculer une approximation de sa dérivée en tout point. La première matrice donne la dérivée horizontale et la seconde donne la dérivée verticale. Si nous définissons A comme étant l'image source, et G_x et G_y les deux images dérivées horizontale et verticale (respectivement) de l'intensité lumineuse de l'image, on peut les calculer ainsi avec l'opération de convolution 2D suivante :

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} * A \quad \text{et} \quad G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix} * A$$

Cet opérateur ainsi défini pouvant se décomposer en un produit d'une moyenne et d'une différentielle, il calcule le gradient lissé.

Par exemple, G_x peut s'écrire :

$$\begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} [-1 \quad 0 \quad 1]$$

Enfin, nous pouvons calculer l'amplitude du gradient via la norme du vecteur ainsi défini en tout point :

$$G = \sqrt{G_x^2 + G_y^2}$$

De même, nous pouvons calculer sa direction :

$$\Theta = \arctan(G_y, G_x)$$

Par exemple, $\Theta = 0$ correspond à un front vertical qui est plus sombre du côté droit.

Filtre Sobel

L'opérateur utilise des matrices de convolution. La matrice (généralement de taille 3x3) subit une convolution avec l'image pour calculer des approximations des dérivées horizontale et verticale. Soit A l'image source, G_x et G_y deux images qui en chaque point contiennent des approximations respectivement de la dérivée horizontale et verticale de chaque point. Ces images sont calculées comme suit :

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * \mathbf{A} \quad \text{et} \quad \mathbf{G}_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * \mathbf{A}$$

En chaque point, les approximations des gradients horizontaux et verticaux peuvent être combinées comme suit pour obtenir une approximation de la norme du gradient :

$$\mathbf{G} = \sqrt{\mathbf{G}_x^2 + \mathbf{G}_y^2}$$

On peut également calculer la direction du gradient comme suit :

$$\Theta = \text{atan2}(\mathbf{G}_y, \mathbf{G}_x)$$

où, par exemple, Θ vaut 0 pour un contour vertical plus foncé à gauche.

Filtre Laplacien

Cette opération de calcul de Laplacien peut alors être appliquée à une image par l'intermédiaire d'un filtrage avec le masque 3*3 sui vant :

$$[0 \ 1 \ 0 ; 1 \ -4 \ 1 ; 0 \ 1 \ 0]$$

D'autres masques peuvent être utilisés

$$[0 \ -1 \ 0 ; -1 \ 4 \ -1 ; 0 \ -1 \ 0] ; [-1 \ -1 \ -1 ; -1 \ 8 \ -1 ; -1 \ -1 \ -1] ; [1 \ -2 \ 1 ; -2 \ 4 \ -2 ; 1 \ -2 \ 1].$$

Il s'agit de détecter les passages par zéro en conservant uniquement les passages les plus marqués. En effet, la technique est particulièrement sensible au bruit en raison de la double dérivation. Il s'agit donc de ne pas considérer le bruit, qui peut très bien se traduire par des oscillations autour de zéro, comme un contour. C'est le rôle du seuil S qui va être utilisé dans cette approche pour ne prendre en compte que les passages par zéro d'amplitude relativement élevée correspondant à des vrais contours de l'image.

Filtre LOG

Le filtre **de Marr et Hildreth**, aussi appelé **LOG** pour **Laplacien Of Gaussien** est très utilisé pour détecter les contours car il permet de résoudre certains problèmes rencontrés avec le filtre laplacien. En effet, le filtre laplacien réalise une double dérivation, qui est extrêmement sensible au bruit.

L'idée est donc d'appliquer le laplacien non pas directement sur l'image, mais sur une image déjà filtrée par un filtre Gaussien (donc potentiellement d'ébruitée). Le filtre LOG dépend donc d'un paramètre σ qui est l'écart-type du filtre Gaussien appliqué en pré-traitement. On choisit σ de façon à réduire l'effet nuisible dû au bruit.

Calcul du noyau : somme des dérivées secondes sur x et y de la gaussienne $G(i,j)$. C'est à dire :

$$Log(i, j) = \frac{(i^2 + j^2 - 2\sigma^2)}{\sigma^4} * G(i, j)$$

Le coefficient central est ajusté de manière à ce que la somme des coefficients du filtre soit égale à zéro.

Méthode de seuillage

Le seuillage est une méthode faisant partie de l'approche basée sur le pixel, même si d'autres références classent cette méthode dans l'approche région de la segmentation, du fait qu'elle met en évidence les régions d'une image.

Un seuil est une valeur numérique correspondant à un paramètre de l'image (exemple : le niveau de gris), et l'histogramme des niveaux de gris est utilisé afin d'en déduire le seuil.

Le seuillage peut être :

-**Global** : un seuil pour toute l'image

-**Local** : un seuil pour une portion de l'image

-**Adaptatif** : un seuil qui s'ajuste selon les images, ou parties de l'image

Comment choisir le seuil ?

- Une valeur obtenue par tests (empirique)
- La valeur moyenne des tons de gris
- La valeur médiane entre le ton maximum et le ton minimum (pour chaque pic)
- Découper l'image en sous-image de manière judicieuse afin d'adapter le seuil à chacune d'entre elles

Quelques méthodes de seuillage adaptatif :

Méthode 1 de seuillage :

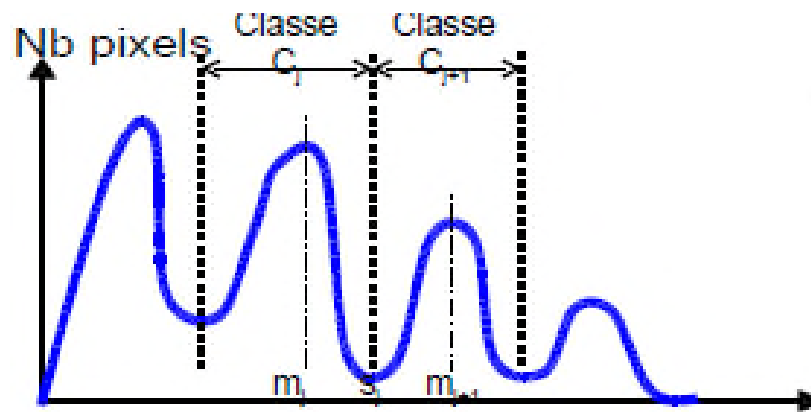
Cette technique est la plus intuitive. On suppose que chaque classe correspond à une gamme distincte de niveaux de gris. L'histogramme est alors m-modal. La position des minima de l'histogramme H permet de fixer les $(m-1)$ seuils nécessaires pour séparer les m classes.

En termes mathématiques, les seuils s_i sont obtenus par

$$H(s_i) = \text{Min} [H(k)] \text{ pour } k \text{ in }]m_i, m_{i+1}[$$

où m_i et m_{i+1} sont les valeurs moyennes (ou les modes) de l'intensité lumineuse dans les classes C_i et C_{i+1} .

Le principe, illustré sur la figure consiste au positionnement des seuils de classification sur ces minima locaux (multi seuillage).



Malgré le développement de techniques robustes visant à faciliter la détection des vallées, cette méthode, bien que simple, est très peu appliquée car les histogrammes traités sont le plus souvent bruités et unimodaux.

Méthode 2 de seuillage :

La procédure est la suivante :

1. On calcule les histogrammes de toutes les composantes de l'image.
2. A partir des histogrammes, on sélectionne le pic le plus significatif sur l'ensemble des histogrammes.

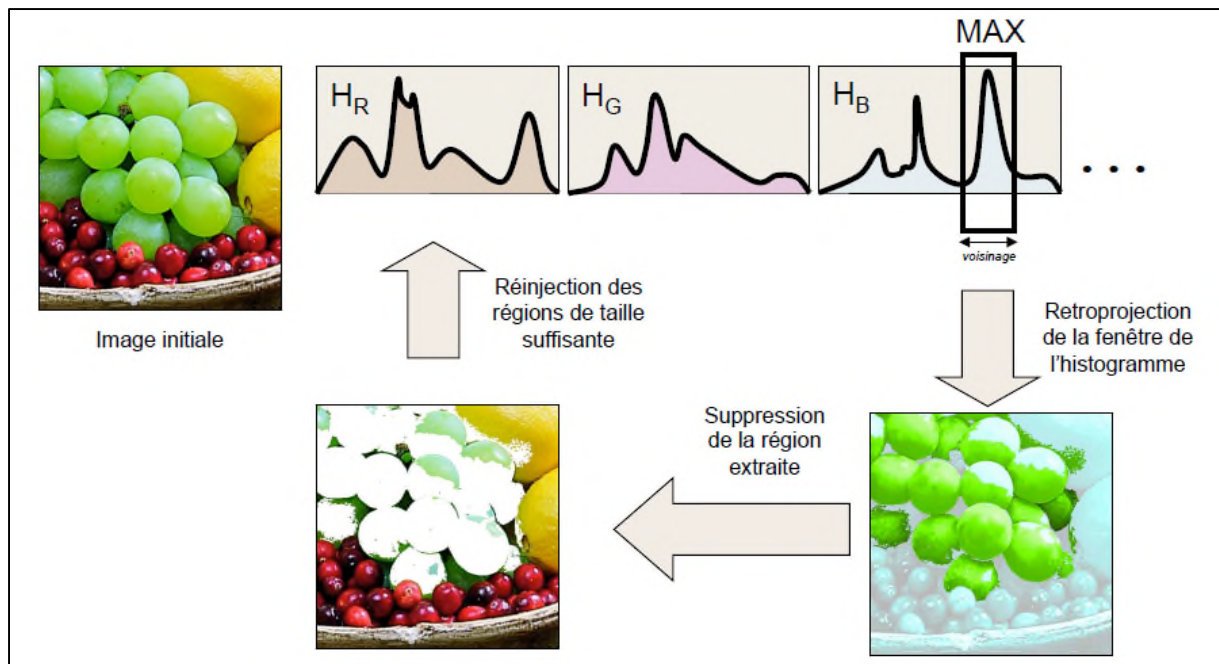
Soit h_x un histogramme. Un pic peut être formellement défini comme une restriction de h_x à un intervalle tel qu'il existe $m \in [n_1, n_2]$ tel que $\forall i \in [n_1, m[, h_x(i) \leq h_x(i + 1)$ et $\forall i \in]m, n_2], h_x(i - 1) \geq h_x(i)$.

Un tel pic peut être caractérisé par sa hauteur $h_x(m)$, sa largeur $(n_2 - n_1)$, sa somme :

$$S = \sum_{i=n_1}^{n_2} h_X(i)$$

L'importance d'un pic peut être caractérisée par sa moyenne (somme sur largeur), son rapport hauteur sur largeur.

3. Une fois le pic sélectionné, les valeurs sont projetées dans l'image, c'est-à-dire qu'on calcule l'image binaire correspondant à l'ensemble des pixels dont la valeur de la composante X est comprise entre n_1 et n_2 .
4. L'ensemble de pixels précédemment formé est retiré du support de l'image. Si le nombre de pixels restant est suffisant, on retourne au (1) en recalculant les histogrammes sur ce nouveau support.



Le principe, illustré sur la figure consiste à sélectionner itérativement la région tonale la plus significative de l'image.

Chaque pixel est décrit selon certains channels : R, G, B => *L'algorithme travaille sur plusieurs histogrammes, un par channel*

Qu'est-ce que la classification automatique ?

La classification automatique est la tâche qui consiste à regrouper, de façon non supervisée (ç-à-d sans l'aide préalable d'un expert), un ensemble d'objets ou plus largement de données, de telle manière que les objets d'un même groupe (appelé cluster) sont plus proches (au sens d'un critère de (dis)similarité choisi) les uns aux autres que celles des autres groupes (clusters). Il s'agit d'une tâche principale dans la fouille exploratoire de données, et une technique d'analyse statistique des données très utilisée dans de nombreux domaines, y compris l'apprentissage automatique, la reconnaissance de formes, le traitement de signal et d'images, la recherche d'information, etc...L'idée est donc de découvrir des groupes au sein des données, de façon automatique.

Quelles données ?

Les données traitées en classification automatique peuvent être des images, signaux, textes, autres types de mesures, etc. Dans le cadre de ce travail les données seront des données multidimensionnelles (une image couleur). Chaque donnée est donc composée de plusieurs variables (descripteurs). Pour le cas de données multidimensionnelles standard, chaque donnée étant de dimension d (donc un point dans l'espace \mathbf{R}^d). Les n données peuvent donc être modélisées comme étant un tableau de n éléments, chaque élément du tableau étant une structure "donnée" comme décrite précédemment. Dans le cas d'une image (couleur), l'image contenant n lignes et m colonnes et donc $n \times m$ pixels couleurs, chaque pixel est composé de trois composantes **RVB**, peut être modélisée par une matrice de 3 ligne et $n \times m$ colonnes.

Qu'est-ce qu'une classe ?

Une classe (ou groupe) est un ensemble de données formée par des données homogènes (qui "se ressemblent" au sens d'un critère de similarité (distance, densité de probabilité, etc)). Par exemple, si on a une base d'image de chiffres manuscrits, on aura la classe des zéros, la classe des uns, etc. De même pour le cas d'images de lettres manuscrites. Une classe peut être aussi une région dans une image couleur, un évènement particulier dans un signal sonore, la classe spam et classes non spam dans le cas de détection de spams dans un mail, etc.

Combien de classes ?

Le nombre de groupes (qu'on notera K), pour commencer, peut être supposé fixe (donné par l'utilisateur). C'est le cas par exemple si l'on s'intéresse à classer des images de chiffres manuscrits (nombre de classes = $10 : 0, \dots, 9$) ou de lettres manuscrites (nombre de classes = nombres de caractères de l'alphabet), etc. Néanmoins, il existe des critères, dits de choix de modèle, qui permettent de choisir le nombre optimal de classes dans un jeu de données. Ce choix est toujours subjectif et fortement dépendant du critère sélectionné pour comparer les résultats.

Quel est le critère d'évaluation de la qualité d'une classification ?

L'analyse visuelle du résultat et des différences est un outil essentiel. Cependant, l'utilisation du système visuel humain en tant qu'outil de jugement de la qualité n'est pas à négliger mais nécessaire pour vérifier la qualité des images obtenues par la classification.

Pour l'évaluer, nous allons utiliser l'analyse visuelle selon la vérité du terrain disponible sur la zone, et nous allons évaluer la classification obtenue au cours de l'exécution par la composition colorée.

Compte-rendu

Dans votre compte-rendu, vous inclurez le *listing* du programme Matlab ou Python ainsi que des commentaires sur les résultats obtenus. Ces commentaires devront mettre en évidence votre sens de l'analyse des résultats. Ils s'appuieront sur les résultats de votre programme, que vous devez donc également inclure dans le compte-rendu.