

Une nouvelle approche pour la classification non supervisée en segmentation d'image

Sébastien Lefèvre

LSIIT – CNRS / Université de Strasbourg
Pôle API, Bd Brant, BP 10413, 67412 Illkirch Cedex
lefevre@lsiit.u-strasbg.fr

Résumé. La segmentation des images en régions est un problème crucial pour l'analyse et la compréhension des images. Parmi les approches existantes pour résoudre ce problème, la classification non supervisée est fréquemment employée lors d'une première étape pour réaliser un partitionnement de l'espace des intensités des pixels (qu'il s'agisse de niveaux de gris, de couleurs ou de réponses spectrales). Puisqu'elle ignore complètement les notions de voisinage des pixels, une seconde étape d'analyse spatiale (étiquetage en composantes connexes par exemple) est ensuite nécessaire pour identifier les régions issues de la segmentation. La non prise en compte de l'information spatiale est une limite majeure de ce type d'approche, ce qui a motivé de nombreux travaux où la classification est couplée à d'autres techniques pour s'affranchir de ce problème. Dans cet article, nous proposons une nouvelle formulation de la classification non supervisée permettant d'effectuer la segmentation des images sans faire appel à des techniques supplémentaires. Plus précisément, nous élaborons une méthode itérative de type k-means où les données à partitionner sont les pixels eux-mêmes (et non plus leurs intensités) et où les distances des points aux centres des classes ne sont plus euclidiennes mais topographiques. La segmentation est alors un processus itératif, et à chaque itération, les classes obtenues peuvent être assimilées à des zones d'influence dans le contexte de la morphologie mathématique. Ce parallèle nous permet de bénéficier des algorithmes efficaces proposés dans ce domaine (tels que ceux basés sur les files d'attente), tout en y ajoutant le caractère itératif des méthodes de classification non supervisée considérées ici. Nous illustrons finalement le potentiel de l'approche proposée par quelques résultats préliminaires de segmentation sur des images artificielles.

1 Introduction

La classification, qu'elle soit supervisée ou non, a toujours été un outil fort employé dans le domaine de l'analyse et du traitement des images numériques, en particulier à des fins de segmentation ou d'interprétation. Dans le même temps, les images peuvent être vues comme des données semi-structurées, complexes, qui offrent de nouvelles perspectives et de nouveaux défis au domaine de la fouille de données et de l'extraction de connaissances.

Nous nous intéressons ici au lien qui existe entre ces deux domaines que sont le traitement des images et la fouille de données, et étudions plus particulièrement le rôle de la classification (non supervisée) pour résoudre un des principaux problèmes du traitement d'image (la segmentation). La segmentation ou partition d'une image en régions peut en effet être obtenue à l'aide d'une classification non supervisée appliquée sur l'ensemble des valeurs des pixels (niveaux de gris, couleurs ou réponses spectrales par exemple) et suivie d'une analyse spatiale (étiquetage en composantes connexes). Quoique très fréquemment employée, cette stratégie de segmentation n'est pas idéale puisque l'information spatiale n'est prise en compte que dans un second temps et complètement ignorée lors de la phase de classification (ce qui a donné lieu à plusieurs solutions ad-hoc pour résoudre ce problème que nous décrivons en section 2.2).

Mais, afin de garantir la prise en compte de l'information spatiale par la classification dans le contexte de la segmentation, pourquoi ne pas appliquer une telle classification directement dans l'espace (spatial) des pixels ? Bien sûr, pour que cette méthode de classification puisse fournir directement un résultat acceptable de segmentation sans faire appel à des traitements supplémentaires, il faut revoir les notions importantes en classification (espace des données, mesure de similarité ou de distance) dans le contexte de la segmentation d'image, ainsi que les algorithmes mis en œuvre. Notre contribution porte ici sur une telle approche, et nous nous intéressons plus spécifiquement à la formulation d'une méthode de classification non supervisée par partitionnement de type k-means dans le contexte de la segmentation d'image.

Le plan de cet article est le suivant. Nous commencerons par rappeler comment la classification non supervisée par partitionnement¹ est employée pour résoudre le problème de la segmentation d'image. Nous expliquerons quelles sont les limites inhérentes à ce type d'approche et quelles ont été les principales solutions proposées pour y remédier. Nous décrivons ensuite plus en détail notre proposition, qui consiste à ne plus effectuer la classification dans l'espace des intensités, mais directement dans l'espace des pixels, en évitant d'utiliser la distance euclidienne et les centres de gravité. Nous ferons le lien avec les travaux en morphologie mathématique, ce qui nous permettra d'exploiter des algorithmes efficaces pour la segmentation d'image basée sur une classification non supervisée (qui peut aussi être considérée comme une segmentation morphologique itérative). Enfin nous illustrerons notre approche par quelques résultats préliminaires de segmentation obtenus sur des images artificielles.

2 Limites de la classification en segmentation d'image

Dans cette section, nous introduirons les notations utilisées ici et expliquerons pourquoi une classification ne permet pas la segmentation d'image. Nous dresserons également un panorama des approches proposées pour adapter les méthodes de classification à la segmentation.

2.1 Segmentation versus classification

Une image est généralement définie comme une fonction $f : \mathcal{E} \rightarrow \mathcal{T}$ qui associe à chaque pixel $p = (x, y)$ de l'espace $\mathcal{E} \subset \mathbb{N}^2$ une valeur v dans \mathcal{T} . Cette valeur peut-être par exemple un niveau de gris ($\mathcal{T} = \{0, \dots, 255\}$), une couleur représentée sous une forme trichromatique ($\mathcal{T} = [0, 1]^3$ ou $\mathcal{T} = \{0, \dots, 255\}^3$) ou encore une réponse spectrale ($\mathcal{T} \subset \mathbb{R}^n$ pour une image à n bandes spectrales).

¹par la suite et dans un souci de concision, nous utiliserons le terme de classification.

La segmentation a pour but de partitionner l'espace des pixels \mathcal{E} d'une image f en un ensemble de K régions $\{R_k\}_{1 \leq k \leq K}$ homogènes au sens d'un certain critère (par exemple les valeurs v des pixels composant chaque région). Il s'agit donc d'une fonction $\pi : \mathcal{E} \rightarrow \mathcal{R}$ qui associe à chaque pixel p l'indice k de la région R_k à laquelle il appartient. Chaque région R_k est construite comme une composante connexe (cf. figure 1(b)), c'est-à-dire un ensemble de pixels adjacents (ou voisins deux à deux) de valeur k . Plus formellement, on définit la notion de chemin discret P_{pq} d'un pixel p à un pixel q comme l'ensemble des pixels $P_{pq} = \{p_i; i = 0, \dots, m\}$ tel que $p_0 = p$, $p_m = q$, et $\forall i = 0, \dots, m-1$, p_i est adjacent à p_{i+1} . Deux pixels $p = (x_p, y_p)$ et $q = (x_q, y_q)$ sont dits λ -adjacents si leur λ -distance vaut 1, soit $d_\lambda(p, q) = 1$. On définit la 4-distance $d_4 = |x_p - x_q| + |y_p - y_q|$, la 8-distance $d_8(p, q) = \max(|x_p - x_q|, |y_p - y_q|)$, et la distance euclidienne $d_E(p, q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2}$. Le voisinage d'un pixel p noté $\mathcal{N}(p)$ est l'ensemble des pixels q adjacents à p . Nous pouvons associer un coût $\omega(P_{pq})$ à chaque chemin, par exemple en dénombrant le nombre de pixels qui le compose, ou en mesurant les distances entre pixels : $\omega(P_{pq}) = \sum_{i=1}^m d(p_{i-1}, p_i)$ avec d une mesure de distance (euclidienne par exemple). Enfin, une région R_k ou composante connexe vérifie $\forall p, q \in R_k, \exists P_{pq}$ avec $\forall p_i \in P_{pq}, p_i \in R_k$ ou autrement dit $\pi(p_i) = k$.

La classification non supervisée aussi appelée partitionnement de données ou *clustering* consiste à regrouper les données au sein d'ensembles homogènes nommés classes. Appliquée à une image, la classification peut être représentée comme une fonction $\pi : \mathcal{E} \rightarrow \mathcal{C}$ qui associe à chaque pixel p l'indice k de la classe C_k à laquelle il appartient. De même que pour les régions R_k , le contenu des classes C_k doit être homogène (par exemple, les pixels qui composent une classe doivent avoir des valeurs v similaires). Cependant, contrairement aux régions R_k produites par la segmentation, les classes C_k n'ont pas la caractéristique d'être des composantes connexes, comme illustré par la figure 1(c).

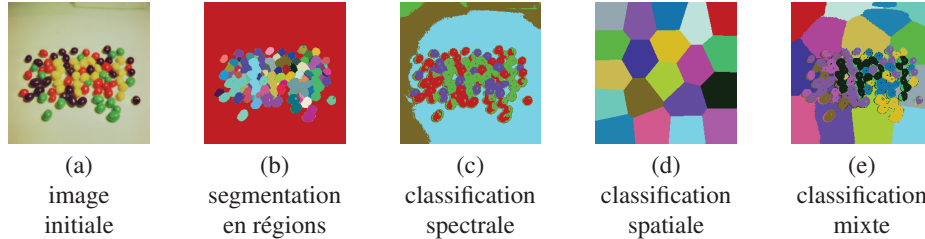


FIG. 1 – Différence entre une segmentation et une classification

Nous nous intéressons plus spécifiquement dans cet article à la classification non supervisée par une méthode de type k-means (Kaufman et Rousseeuw, 1990). Dans cet algorithme, le nombre K de classes $\{C_k\}_{1 \leq k \leq K}$ est défini a priori (même s'il existe d'autres méthodes de partitionnement s'affranchissant de cette condition). Chaque classe C_k est caractérisée par son centre noté c_k , nous noterons également s_k le nombre d'éléments de la classe C_k . L'initialisation des centres des classes peut être guidée par des hypothèses disponibles sur les données, ou alors réalisée de façon aléatoire en l'absence de connaissance sur les données. L'algorithme procède ensuite en deux étapes itératives effectuées jusqu'à convergence : dans un premier temps, on calcule pour chaque pixel une distance aux différents centres des classes et on af-

Une nouvelle approche pour la classification non supervisée en segmentation d'image

fecte le pixel à la classe dont le centre est le plus proche ; ensuite, on met à jour les centres des classes en s'appuyant sur la nouvelle partition des données.

L'application de cet algorithme à la segmentation d'images s'effectue en utilisant non pas les pixels p directement mais leurs valeurs dans l'image $f(p)$. L'algorithme k-means dans sa formulation originale va chercher à minimiser une fonction de coût global définie par :

$$J = \sum_{k=1}^K \sum_{p \in C_k} (f(p) - c_k)^2 \quad (1)$$

et peut être décrit par les étapes suivantes (où les itérations sont représentées par une notation en exposant) :

1. initialisation aléatoire de chaque centre c_k^0 dans l'espace \mathcal{T} ;
2. pour chaque pixel p , calcul de la distance $d(f(p), c_k^l)$ aux différents centres des classes c_k^l , et affectation à la classe la plus proche $\pi^l(p) = \arg \min_k d(f(p), c_k^l)$;
3. mise à jour des centres c_k des classes en s'appuyant sur la nouvelle partition π^l pour recalculer les centres de gravité, soit $c_k^{l+1} = \text{moy}\{f(p_i); \pi^l(p_i) = k\}$ avec moy la fonction moyenne ;
4. arrêt si $\pi^l = \pi^{l-1}$, sinon retour à l'étape 2.

La distance d mise en œuvre ici est la distance euclidienne calculée dans l'espace \mathcal{T}^n , soit $d_E(v, w) = \sqrt{\sum_{i=1}^n (v_i - w_i)^2}$ avec $v = (v_1, \dots, v_n)$ et $w = (w_1, \dots, w_n)$. Le critère de convergence (stabilité de la partition) peut être quelque peu relâché et remplacé par une convergence de la fonction de coût, voire un nombre fini d'itérations.

Dans l'algorithme précédent (cf. également figure 1(c)), il est clair que la position du pixel p (ses coordonnées (x, y)) a été complètement ignorée, au profit de sa valeur $f(p)$. Deux pixels p et q aux valeurs proches seront donc très probablement associés à une même classe C_j même s'ils forment deux composantes connexes disjointes dans l'image des classes π . Pour obtenir une segmentation, c'est-à-dire une partition de l'image en zones connexes homogènes, des traitements supplémentaires sont nécessaires.

2.2 Classification tenant compte de l'information spatiale

L'obtention d'une segmentation à partir d'une classification nécessite la prise en compte de l'information spatiale (Haralick et Shapiro, 1985). Nous dressons ici un bref panorama des principales techniques (non exclusives) proposées dans la littérature. Dans ce panorama, nous ne tenons pas compte des approches par champs de Markov qui combinent classification et régularisation spatiale, même si elles peuvent être assimilées à des généralisations des méthodes de classification telles que les k-means (Pappas, 1992). Nous nous intéresserons plutôt ici aux approches telles que k-means ou fuzzy c-means.

La solution la plus fréquemment rencontrée est un post-traitement : un étiquetage en composantes connexes est appliqué suite à l'étape de classification. Il permet de regrouper au sein d'une même région R_k tous les pixels p adjacents qui ont été affectés à la même classe C_j . En reprenant la formulation introduite précédemment, une région R_k est alors une composante connexe vérifiant $\forall p, q \in R_k, \exists P_{pq}$ avec $\forall p_i \in P_{pq}, p_i \in C_j$. L'information spatiale peut aussi être exploitée via une segmentation initiale en régions, dont le résultat (les régions) fera l'objet d'une classification postérieure.

Cependant, appliquer une classification au niveau des pixels puis regrouper les pixels de même classe en composantes connexes produit généralement une carte de segmentation où de nombreuses régions ne consistent qu'en des pixels isolés (et non des agrégats de pixels). Une étape de filtrage peut donc être intégrée au processus pour supprimer ces pixels isolés et les réaffecter aux régions voisines. Dans le cas d'une classification floue, le post-traitement à appliquer peut être lui-même un algorithme de segmentation tel que la croissance de régions (Eum et al., 1996), la fusion de régions (Chen et Lu, 2002), ou encore la méthode de ligne de partage des eaux avec marqueurs (Lezoray et Cardot, 2002).

L'intégration d'informations spatiales peut également être considérée comme une étape de pré-traitement. Ainsi, on peut décrire chaque pixel p par sa valeur v mais également par ses coordonnées (x, y) (Krishnapuram et Freg, 1992) comme illustré en figure 1(e). Il est aussi possible d'effectuer une interpolation de l'image, où la valeur estimée $\hat{f}(p)$ en chaque pixel p est calculée à partir de ses 4 ou 8 voisins (sous la forme d'une moyenne ou d'une médiane par exemple). La classification s'applique alors sur les pixels en considérant les valeurs \hat{f} seules ou en complément de l'image initiale f (Turi, 2001; Chen et Zhang, 2004).

Il est possible de tenir compte de l'information spatiale directement au sein de l'algorithme de classification. Ilea et Whelan (2008) suggère de coupler l'information de texture (mesurée par un gradient) à une information de couleur obtenue par application d'un k-means sur une image lissée par filtrage anisotropique et représentée dans deux espaces couleur différents. La mesure de dissimilarité d'un pixel à une classe s'appuie sur ces informations, qui sont associées aux pixels mais aussi aux classes, et mises à jour à chaque itération pour suivre l'évolution du contenu des classes.

Liew et al. (2000) propose de définir dans l'algorithme fuzzy c-means la dissimilarité entre un pixel et le centre d'une classe en tenant compte de l'ensemble des 8 voisins du pixel : la contribution de chaque voisin (calculée comme une distance au centre de la classe concernée) sera proportionnelle à la similarité du voisin avec le pixel considéré, la similarité étant mesurée dans l'espace des caractéristiques. Cette approche est inspirée de (Tolias et Panas, 1998) où la similarité d'un voisin au pixel à classer modifiait la probabilité d'appartenance de ce pixel aux classes (ajout ou suppression d'une constante). Cette similarité entre voisins peut également être calculée directement dans l'espace des appartenances aux classes (Noordam et al., 2000; Xia et al., 2007). Pham (2001) suggère une approche similaire en utilisant une classification probabiliste basée sur des indicateurs de krigeage des voisins.

Appliquer un classifieur dans un cadre multi-échelle, à l'aide d'une méthode hiérarchique, offre un autre moyen de prendre en compte le contexte spatial. Dans (Luo et al., 2003), les auteurs proposent un nouvel algorithme où sont successivement appliquées les étapes suivantes : sous-échantillonnage de l'image, calcul des caractéristiques liées à la couleur et la texture, classification contrainte spatialement, et fusion de régions. Ces deux dernières étapes sont effectuées itérativement à chaque niveau de la hiérarchie, permettant ainsi d'isoler les principales composantes connexes (et d'agréger les pixels restants à ces composantes), puis de réappliquer l'algorithme de k-means sur chaque composante connexe indépendamment des autres. Ce principe de classifications successives est également rencontré dans les travaux de (Cheng et Sun, 2000) où les clusters sont construits uniquement à partir des pixels de voisinage homogène. Enfin, Leydier et al. (2004) proposent d'appliquer un k-means sur une fenêtre glissante dans l'image où les centres des classes sont propagés lors du parcours de l'image.

Pour résumer, les solutions pour la spatialisation des méthodes de classification afin de les

Une nouvelle approche pour la classification non supervisée en segmentation d'image

rendre opérationnelles pour la segmentation des images sont très variées : pré-traitement, post-traitement, ajout d'attributs, approche multi-échelle, modification de la classification d'un pixel en fonction de ses voisins, etc. Cependant, à la connaissance de l'auteur, l'application d'une classification directement dans le domaine spatial n'a pas encore été suggérée, alors qu'il s'agit d'une solution très intuitive pour prendre en compte le caractère spatial d'une segmentation. Nous proposons donc dans la section suivante une telle solution, qui peut être vue comme une classification spatiale par un algorithme de k-means.

3 Un autre usage de la classification en segmentation d'image

Pour appliquer avec succès une méthode de classification au problème de la segmentation d'image, il est nécessaire de prendre en compte le contexte spatial. A l'intégration d'étapes supplémentaires dans la classification, nous préférons la reformulation de la classification dans un contexte spatial. Ainsi nous proposons de spatialiser la méthode de classification afin que celle-ci puisse être utilisée pour effectuer la segmentation des images.

3.1 Spatialisation de l'algorithme de classification

Nous avons vu dans la section 2.1 que la seule différence entre une segmentation et une classification était l'absence de contrainte de connexité spatiale dans les résultats de la seconde. Modifier le comportement d'un algorithme de classification pour qu'il garantisse la connexité spatiale des classes produites (ou régions dans ce contexte) permettrait alors de disposer d'une méthode de classification intrinsèquement capable de résoudre le problème de la segmentation d'image. Les algorithmes de classification étant par essence très variés, nous avons fait le choix de nous intéresser ici à l'un des exemples les plus connus, l'algorithme des k-means, et d'étudier comment nous pouvions garantir à cet algorithme la propriété de connexité spatiale. Autrement dit, notre travail porte sur la spatialisation de l'algorithme k-means.

L'information spatiale est directement présente dans les images au travers des notions de connexité et d'adjacence entre pixels voisins. Aussi nous proposons d'appliquer la classification dans l'espace \mathcal{E} des pixels et non plus un espace de caractéristiques (par exemple l'espace \mathcal{T} des valeurs des pixels). Les différents calculs (distance d entre points et centres des classes lors de l'étape 2, mise à jour des centres des classes avec la fonction moyenne moy lors de l'étape 3) s'effectuent également dans l'espace \mathcal{E} , en considérant p et non plus $f(p)$ dans les différentes formules. De même les centres des classes c_k seront définis dans \mathcal{E} (i.e. ce sont des pixels de l'image) et non plus dans \mathcal{T} .

Cependant, l'application directe de l'algorithme k-means dans l'espace \mathcal{E} n'aura pas le résultat escompté pour la segmentation : puisque les pixels de l'image sont disposés régulièrement sur une grille carrée de taille finie, la non utilisation des valeurs des pixels lors de l'affectation des pixels aux classes aura pour effet de fournir un résultat indépendant de l'image analysée (figure 1(d)) ! Une solution intermédiaire doit être mise en oeuvre pour tenir compte à la fois des pixels p et de leurs valeurs $f(p)$ lors du calcul de la distance $d(p, c_k)$ entre pixels et centres des classes. Elle fera l'objet de la section suivante.

Un autre problème doit également être résolu pour garantir un comportement adéquat de l'algorithme des k-means dans le contexte de la segmentation. Il s'agit de l'incapacité de l'algorithme à traiter correctement des ensembles non convexes. Or lors d'une segmentation en

régions, on ne peut pas faire l'hypothèse que toutes les régions soient convexes. Dans le cas d'une région concave, la mise à jour du centre de la classe (ou de la région) par le calcul du centre de gravité peut aboutir au positionnement du centre en dehors de la classe. Comme nous le verrons dans la section suivante, nous nous basons sur l'hypothèse d'appartenance du centre à la région ou classe pour garantir la connexité spatiale des classes produites. Il nous faut donc remplacer le centre de gravité par une autre mesure et nous proposons d'utiliser une médiane (aboutissant ainsi à un algorithme de type k-médians plutôt que k-means). De plus, nous souhaitons conserver les centres des classes dans \mathcal{E} , c'est-à-dire choisir les centres des classes parmi les pixels présents dans l'image, à la façon d'un k-prototypes (Han et al., 2001) (tout en conservant le schéma directeur d'un k-means). Pour obtenir une telle mesure médiane dans \mathcal{E} , nous proposons de définir le centre d'une classe comme son pixel central, que l'on peut obtenir par des méthodes de morphologie mathématique (Soille, 2003) : érosion ultime de la région, maximum de la transformée de distance à l'extérieur de la région, etc.

Notons qu'ici nous avons fait le choix de considérer indépendamment les espaces \mathcal{E} des pixels et \mathcal{T} des valeurs. Nous aurions pu aussi formuler une solution directement dans l'espace $\mathcal{E} \times \mathcal{T}$ en imposant certaines contraintes quant aux classes à obtenir, comme dans le cas des nuées dynamiques (Diday, 1971) ou des k-medoids (Kaufman et Rousseeuw, 1990). Cependant, appliquer une classification dans l'espace $\mathcal{E} \times \mathcal{T}$ se traduirait par une complexité algorithmique plus importante.

3.2 Distance topographique et garantie de connexité

Pour tenir compte à la fois de ses coordonnées et de sa valeur lors de l'affectation d'un pixel aux classes, nous proposons de combiner p et $f(p)$ au travers d'une distance topographique dont nous donnons quelques définitions ci-dessous. Notons que la mise à jour des centres des classes peut elle aussi tenir compte de cette distance et des valeurs p et $f(p)$, comme nous le verrons par la suite.

La représentation d'une image comme une surface topographique associe à chaque pixel p ou position une altitude, c'est-à-dire la valeur $f(p)$ du pixel. Dans ce contexte, la notion de coût d'un chemin P_{pq} peut être reformulée en tenant compte de sa topographie ou des dénivellés rencontrés pour aller de p à q . Ainsi, plusieurs définitions du coût topographique d'un chemin ont été proposées, notamment celles de Prêteux (1992), Meyer (1994) et Philipp-Foliguet (2000). La notion de coût introduite par Prêteux peut être formulée simplement dans le cas d'une image discrète et finie comme le pixel le plus élevé sur le chemin :

$$\omega(P_{pq}) = \max_{p_i \in P_{pq}} f(p_i) \quad (2)$$

ou autrement dit la crête du chemin, avec la convention $\omega(P_{pp}) = -\infty$.

Meyer utilise quant à lui la notion de pente maximale définie pour un pixel p comme

$$\delta(p) = \max_{\substack{q \in \mathcal{N}(p) \\ f(q) < f(p)}} \left(\frac{f(p) - f(q)}{d(p, q)} \right) \quad (3)$$

Une nouvelle approche pour la classification non supervisée en segmentation d'image

qui peut être calculée simplement par une érosion morphologique, et défini le coût topographique entre deux pixels voisins de la façon suivante :

$$q \in \mathcal{N}(p), \quad \omega(p, q) = \begin{cases} \delta(p) & \text{si } f(p) > f(q) \\ (\delta(p) + \delta(q))/2 & \text{si } f(p) = f(q) \\ \delta(q) & \text{si } f(p) < f(q) \end{cases} \quad (4)$$

Le coût topographique d'un chemin est alors défini par

$$\omega(P_{pq}) = \sum_{i=1}^m (d(p_{i-1}, p_i) \times \omega(p_{i-1}, p_i)), \quad P_{pq} = \{p_i\}_{1 \leq i \leq m} \quad (5)$$

Philipp-Foliguet propose une définition plus simple qui évite la recherche de la crête ou des pentes maximales, utilisant un coefficient de pondération κ (généralement $\kappa = 1$) :

$$\omega(P_{pq}) = \sum_{i=1}^m (\kappa d(p_{i-1}, p_i) + |f(p_{i-1}) - f(p_i)|), \quad P_{pq} = \{p_i\}_{1 \leq i \leq m} \quad (6)$$

Finalement, on peut définir la distance topographique $d_T(p, q)$ comme le coût minimum d'un chemin reliant p à q , soit $d_T(p, q) = \min(\omega(P_{pq})) \forall P_{pq}$. Notons que d_T est une distance au vrai sens du terme uniquement dans la définition de Philipp-Foliguet. Dans les autres cas, la propriété de séparation n'est pas respectée puisque $\forall p, q, d_T(p, q) = 0 \Leftrightarrow p = q$. Prêteux propose d'utiliser $\exp(d_T(p, q))$ tandis que Meyer présente différentes techniques de modification de la surface topographique pour traiter le cas problématique des plateaux (où on a $d_T(p, q) = 0$ avec $p \neq q$). Remarquons qu'une autre mesure de pseudo-distance (puisque ne respectant pas la propriété de séparation) pourrait être obtenue en définissant simplement le coût comme $\omega(P_{pq}) = \sum_{i=1}^m |f(p_i) - f(p_{i-1})|$. Par la suite nous utiliserons une définition simplifiée de la mesure de Philipp-Foliguet en considérant une 8-distance, où le coût topographique entre deux pixels voisins est $\omega(p, q) = |f(p) - f(q)| + 1, q \in \mathcal{N}_8(p)$.

Nous rappelons que notre objectif est de garantir la connexité spatiale des classes C_k créées par l'algorithme de classification. Vérifions que cette propriété est assurée par l'étape d'affectation d'un pixel à une classe. En reprenant la définition introduite en section 2.1, une classe C_k est connexe si $\forall p, q \in C_k, \exists P_{pq}$ tel que $\forall p_i \in P_{pq}, p_i \in C_k$. Pour simplifier, nous utiliserons dans ce cas la notation $P_{pq} \in C_k$. L'affectation d'un pixel p à C_k s'effectue si $d_T(p, c_k) < d_T(p, c_j), \forall j \neq k$. Naturellement le centre c_k de la classe C_k restera affecté à C_k puisque $d_T(c_k, c_j) > d_T(c_k, c_k) = 0, \forall j \neq k$. La notion de connexité peut alors s'écrire comme $\forall p, q \in C_k, \exists (P_{pc_k}, P_{qc_k}) \in C_k$ ou plus simplement comme $\forall p \in C_k, \exists P_{c_k p} \in C_k$. Soit q le pixel précédant p dans ce chemin, on a donc $q = \arg \min_{p_i \in \mathcal{N}(p)} d_T(c_k, p_i)$ et $d_T(c_k, p) = d_T(c_k, q) + d_T(q, p)$. On peut prouver la connexité de C_k en montrant que l'hypothèse $q \notin C_k$ n'est pas valide. Soit $q \in C_j, j \neq k$, on a alors $d_T(c_k, q) > d_T(c_j, q)$. En combinant les formules précédentes, on obtient $d_T(c_k, p) > d_T(c_j, q) + d_T(q, p) \geq d_T(c_j, p)$. Cette inéquation n'est pas valide puisqu'elle signifierait l'affectation de p à C_j . Le schéma proposé pour l'affectation des pixels aux classes permet donc d'assurer la propriété de connexité spatiale des classes construites.

4 Implémentation et expérimentation

L'algorithme de k -means² que nous proposons ici diffère de la version originale uniquement sur les aspects suivants : (1) l'espace dans lequel sont représentés les pixels et choisis les centres (\mathcal{E} et non plus \mathcal{T}) ; (2) la mesure de distance $d_T(p, c_k)$ qui n'est plus euclidienne mais topographique ; (3) la méthode du calcul du centre des classes qui n'utilise plus la moyenne mais la médiane. Ces changements permettent à l'algorithme de garantir un comportement adéquat pour la segmentation d'image. De plus, ils rapprochent notre solution du cadre de la morphologie mathématique, ce qui va nous permettre de disposer d'implémentations efficaces.

En effet, la distance topographique a été utilisée dans le contexte de la morphologie mathématique pour définir des méthodes de segmentation telles que la ligne de partage des eaux (LPE) ou le squelette par zones d'influence (SKIZ) (Prêteux, 1992; Meyer, 1994). L'étape d'affectation des pixels aux classes de notre algorithme k -means s'assimile donc à une segmentation morphologique à partir de marqueurs qui correspondent aux centres des classes. Il est alors possible de tirer parti des développements algorithmiques en morphologie mathématique, et notamment des algorithmes par file d'attente hiérarchique pour la segmentation guidée par des marqueurs dont on trouvera des exemples d'implémentation dans l'ouvrage de Soille (2003).

Comme nous l'avons indiqué, la mise à jour des centres des classes peut également s'effectuer à l'aide de méthodes morphologiques. Ainsi, on peut déterminer le centre c_k d'une classe C_k à partir de l'érodé ultime de C_k , c'est-à-dire les derniers pixels présents dans C_k après des érosions morphologiques successives. Nous avons fait ici le choix plus logique d'utiliser le maximum de la transformée de distance des pixels de C_k au fond défini par $\mathcal{E} \setminus C_k$ (le fond pouvant être limité judicieusement aux bords extérieurs de la classe C_k). Ce centre est défini par $c_k = \arg \max_{p \in \mathcal{E}} d(p, \mathcal{E} \setminus C_k)$ avec $d(p, X) = \min\{d(p, x); x \in X\}$ la distance de p à l'ensemble X . La distance $d(p, x)$ peut être la distance euclidienne d_E ou de préférence la distance topographique d_T , ce qui permet ainsi d'assurer une meilleure stabilité à l'algorithme puisque c'est la même mesure de distance qui est utilisée lors de l'étape d'affectation. La cohérence de ces deux étapes garantit la convergence de l'algorithme via l'optimisation du critère de qualité classique du k means (1) en considérant les distances topographiques.

Notre méthode peut donc également être assimilée à une segmentation morphologique itérative. L'utilisation d'algorithmes morphologiques efficaces par file d'attente hiérarchique nous permet d'assurer une complexité relativement faible à l'algorithme, indépendante du nombre de classes K , et de l'ordre de $O(P)$ par itération, avec P le nombre de pixels. Plus précisément, lors de l'étape d'affectation, nous ne calculons pas de façon exhaustive les K transformées de distances puisqu'il suffit de connaître en chaque pixel le centre le plus proche. En propageant les distances depuis les centres des classes, on peut traiter et classer itérativement les pixels de distance croissante insérés dans la file (où seule la première occurrence du pixel est considérée). L'actualisation des centres s'effectue selon le même principe et nécessite l'application de K transformées de distance incomplètes (chacune étant limitée aux pixels de la classe concernée), ce qui revient à ne traiter qu'une fois chaque pixel. Les complexités de ces deux étapes sont donc linéaires en $O(P)$. A titre de comparaison, nous rappelons que l'algorithme k -means à un coût par itération de $O(KP)$, K pouvant être relativement grand (tout en étant inférieur à P) puisqu'il s'agit ici du nombre d'objets présents dans l'image et non plus du nombre de classes.

²Notons que l'appellation k -means est ici un abus de langage et que notre algorithme est plutôt de type k -means.

Une nouvelle approche pour la classification non supervisée en segmentation d'image

Nous avons réalisé quelques expérimentations préliminaires de notre algorithme pour valider son comportement. Nous présentons ici le résultat d'une segmentation en 3 régions d'une image artificielle (figure 2). Nous pouvons constater sur cette image que la robustesse au bruit de l'algorithme semble satisfaisante, contrairement à la méthode originale des k-means. Bien évidemment, il ne s'agit ici que de résultats préliminaires et des expérimentations à plus grande échelle doivent être menées pour démontrer l'intérêt effectif des idées proposées.

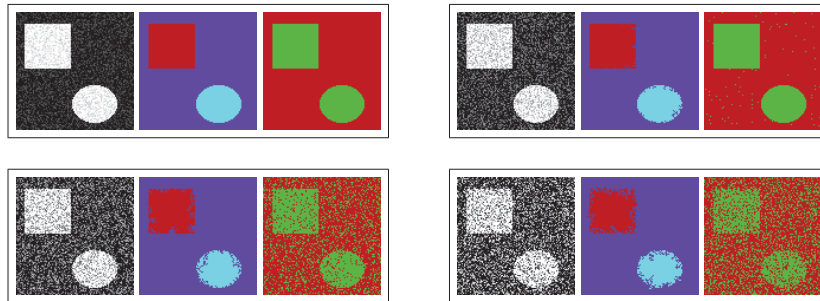


FIG. 2 – Résultats sur images artificielles : (pour chaque encadré, de gauche à droite) image originale bruitée, classification proposée avec 3 régions, k-means des intensités avec 2 classes.

5 Conclusion

La classification est une des solutions fréquemment employées pour réaliser la segmentation d'une image en régions, qui permet de passer d'une représentation sous forme de pixels à des objets pouvant par la suite être analysés afin d'interpréter le contenu des images. Cependant, les méthodes de classification sont le plus souvent appliquées sur l'espace des valeurs des pixels, sans tenir compte de l'information spatiale pourtant primordiale dans le contexte de la segmentation. Pour résoudre ce problème et ainsi tenir compte du contexte spatial de chaque pixel lors de la classification, les nombreuses solutions proposées dans la littérature ont en commun l'intégration d'étapes supplémentaires lors de la classification.

Dans cet article, nous proposons une approche radicalement différente qui consiste à appliquer la classification dans le domaine spatial (des pixels) plutôt que dans le domaine spectral (des valeurs des pixels). Ainsi nous pouvons conserver la méthode de classification en l'état (avec quelques ajustements liés à l'application spatiale de la méthode) et ne pas intégrer d'étapes supplémentaires ad hoc. Nous illustrons notre propos par la spatialisation de la méthode k-means, en remplaçant la distance euclidienne par une distance topographique et en modification le calcul des centres des classes. En outre, un parallèle avec les travaux en morphologie mathématique permet de disposer d'algorithmes efficaces pour la segmentation.

Cette nouvelle approche pour la classification non supervisée en segmentation d'image s'accompagne de nombreuses perspectives. Parmi les aspects à étudier en premier lieu, il est nécessaire de s'affranchir des problèmes relatifs à l'algorithme des k-means tels que le choix des centres initiaux ou la connaissance du nombre de classes a priori. Nous souhaitons bien évidemment évaluer notre méthode dans le cas d'images naturelles, et les premières expérimentations menées dans ce cadre nous ont indiqué que même si l'algorithme convergeait

rapidement, la mesure de distance choisie ici introduisait des frontières artificielles. Il est également possible d'envisager l'affectation floue d'un pixel à une classe (Philipp-Foliguet, 2000) et la prise en compte de contraintes spatiales dans la classification (Han et al., 2001), qui dans le contexte de la segmentation permettrait de tenir compte des zones de fort gradient représentatives des frontières entre objets.

Références

- Chen, S. et D. Zhang (2004). Robust image segmentation using fcm with spatial constraints based on new kernel-induced distance measure. *IEEE Transactions on Systems, Man, and Cybernetics – Part B : Cybernetics* 34(4), 1907–1916.
- Chen, T. et Y. Lu (2002). Color image segmentation – an innovative approach. *Pattern Recognition* 35(2), 395–405.
- Cheng, H. et Y. Sun (2000). A hierarchical approach to color image segmentation using homogeneity. *IEEE Transactions on Image Processing* 9(12), 2071–2082.
- Diday, E. (1971). Une nouvelle méthode en classification automatique et reconnaissance des formes : la méthode des nuées dynamiques. *Revue de Statistique Appliquée* 19(2), 19–33.
- Eum, K., J. Lee, et A. Venetsanopoulos (1996). Color image segmentation using a possibilistic approach. In *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1150–1155.
- Han, J., M. Kamber, et A. Tung (2001). Spatial clustering methods in data mining : a survey. In H. Miller et J. Han (Eds.), *Geographic Data Mining and Knowledge Discovery*. Taylor & Francis.
- Haralick, R. et L. Shapiro (1985). Image segmentation techniques. *Computer Vision, Graphics and Image Processing* 29, 100–132.
- Ilea, D. et P. Whelan (2008). Ctex – an adaptive unsupervised segmentation algorithm based on color-texture coherence. *IEEE Transactions on Image Processing* 17(10), 1926–1939.
- Kaufman, L. et P. Rousseeuw (1990). *Finding groups in data : an introduction to cluster analysis*. New York : Wiley.
- Krishnapuram, R. et C. Fleg (1992). Fitting an unknown number of lines and planes to image data through compatible cluster merging. *Pattern Recognition* (25), 385–400.
- Leydier, Y., F. L. Bourgeois, et H. Emptoz (2004). Sérialisation du k-means pour la segmentation des images en couleur : Application aux images de documents et autres. In *Colloque International Francophone sur l'Écrit et le Document (CIFED)*.
- Lezoray, O. et H. Cardot (2002). Cooperation of color pixel classification schemes and color watershed : a study for microscopical images. *IEEE Transactions on Image Processing* 11(7), 783–789.
- Liew, A., S. Leung, et W. Lau (2000). Fuzzy image clustering incorporating spatial continuity. *IEE Proceedings on Vision, Image and Signal Processing* 147(2), 185–192.
- Luo, M., Y. Ma, et H. Zhang (2003). A spatial constrained k-means approach to image segmentation. In *Pacific Rim Conference on Multimedia*, pp. 738–742.

- Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing* 38, 113–125.
- Noordam, J., W. van den Broek, et L. Buydens (2000). Geometrically guided fuzzy c-means clustering for multivariate image segmentation. In *IAPR International Conference on Pattern Recognition (ICPR)*, pp. 462–465.
- Pappas, T. (1992). An adaptive clustering algorithm for image segmentation. *IEEE Transactions on Signal Processing* 40, 901–914.
- Pham, T. (2001). Image segmentation using probabilistic fuzzy c-means clustering. In *IEEE International Conference on Image Processing (ICIP)*, pp. 722–725.
- Philipp-Foliguet, S. (2000). Segmentation d'images en régions floues. In *Rencontres Franco-phones sur la Logique Floue et ses Applications*, pp. 186–196.
- Prêteux, F. (1992). Watershed and skeleton by influence zones : a distance-based approach. *Journal of Mathematical Imaging and Vision* 1, 239–255.
- Soille, P. (2003). *Morphological Image Analysis : Principles and Applications*. Berlin : Springer-Verlag.
- Tolias, Y. et S. Panas (1998). Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions. *IEEE Transactions on Systems, Man, and Cybernetics – Part A : Systems and Humans* 28(3), 359–369.
- Turi, R. (2001). *Clustering-Based Colour Image Segmentation*. Thèse de doctorat, Monash University. 373 pages.
- Xia, Y., D. Feng, T. Wang, R. Zhao, et Y. Zhang (2007). Image segmentation by clustering of spatial patterns. *Pattern Recognition Letters* 28, 1548–1555.

Summary

Image segmentation is a fundamental problem in image analysis and understanding. Among the existing approaches proposed to solve this problem, unsupervised classification is often involved in a first step to partition the space of pixel intensities (i.e. either grey levels, colours or spectral signatures). Since it completely ignores pixel neighbourhoods, a second step is then necessary to ensure spatial analysis (e.g. with a connected component labeling) in order to identify the regions built from the segmentation process. The lack of spatial information is a major drawback of this kind of approaches, thus many solutions (where classification is used together with other techniques) have been proposed in the literature.

In this paper, we propose a new formulation of the unsupervised classification which is able to perform image segmentation without requiring the need for some additional techniques. More precisely, we introduce a kmeans-related method where data to be clustered are pixels themselves (and not anymore their intensities) and where distances between points and class centers are not anymore euclidean but topographical. Segmentation is then an iterative process, where at each iteration resulting classes can be seen as influence zones in the context of mathematical morphology. This comparison provide us some efficient algorithms proposed in this field (such as hierarchical queue-based solutions), while adding the iterative property of unsupervised classification methods considered here. Finally, we illustrate the potential of our approach by some preliminary segmentation results obtained on artificial images.