

Exploitation de connaissances structurelles en classification d'images : utilisation de méthodes heuristiques d'optimisation combinatoire

Hiroataka SUZUKI ¹, Pascal MATSAKIS ², Jacky DESACHY ³

¹ Université Paul Sabatier, IRIT, 118 route de Narbonne, 31062 Toulouse Cedex, France
Hiroataka.Suzuki@irit.fr

² University of Missouri-Columbia, Computer Engineering & Computer Science Dept.,
201 Engineering Building West, Columbia, MO 65211, USA
pmatsakis@ece.missouri.edu

³ Université des Antilles et de la Guyane, UFR Sciences Exactes et Naturelles,
Campus de Fouillole, 97159 Pointe à Pitre (Guadeloupe)
Jacky.Desachy@univ-ag.fr

Résumé

Cet article traite de l'intégration de connaissances structurelles dans le processus de classification d'une image satellite. Supposons que nous ayons mis en évidence dans une image, par un moyen quelconque, un certain nombre d'objets. Supposons d'autre part que nous ayons récolté auprès d'un expert un certain nombre de connaissances structurelles sur les classes auxquelles nous sommes tentés de rattacher ces objets. Se posent alors trois questions fondamentales. Comment représenter les connaissances ? Comment mesurer l'adéquation entre les objets et les connaissances censées les concerner ? Comment utiliser de telles mesures pour mieux classifier l'image ? Nous décrivons dans cet article une approche fondée sur la représentation des connaissances par un système d'inférence floue ; le calcul, dans une étape préliminaire, d'une partition floue et de la partition nette qui lui est associée ; l'utilisation de méthodes heuristiques d'optimisation combinatoire. Des résultats expérimentaux obtenus à partir de données et connaissances de synthèse ainsi que de l'image RVB d'un objet simple très structuré montrent le bien-fondé de notre approche.

Mots-clés

Classification d'images, connaissances structurelles, logique floue, optimisation combinatoire, recuit simulé, méthode tabou, algorithme génétique.

1. Introduction

Cette recherche s'inscrit dans le domaine de la classification d'images satellites. Les systèmes de classification automatique actuels sont encore loin de pouvoir rivaliser avec le photo-interprète humain. De considérables efforts ont pourtant été entrepris dans la conception de tels systèmes. D'innombrables classifieurs fondés sur l'analyse spectrale des pixels ont été proposés. Mais l'information spectrale n'est pas suffisante, et, globalement, les résultats obtenus ne sont pas satisfaisants. Il est nécessaire d'exploiter des informations supplémentaires, provenant de sources diverses (photographies aériennes, modèle numérique de terrain, carte géologique...). C'est le principe de la classification multisource [1]. D'autre part, les pixels ne peuvent pas être traités

indépendamment les uns des autres sans que de sérieuses limitations n'apparaissent. En particulier, lors du classement d'un pixel, il semble naturel de se soucier des voisins immédiats du pixel en question. Autrement dit, il semble naturel d'utiliser des informations contextuelles [2]. La théorie des sous-ensembles flous [3] offre un cadre bien adapté à la conception de classifieurs multisources et contextuels [4][5][6][7][8]. Cependant, dans de nombreux cas, le photo-interprète humain fait implicitement appel à des connaissances *structurelles*, qui ne sont pas exploitées par les systèmes précédents. L'homme ne se limite pas à la considération d'informations contextuelles de voisinage. Il s'attache aussi à la forme des régions qu'il perçoit dans l'image, et aux relations spatiales entre ces régions. C'est cette question que nous abordons ici. Supposons donc que nous ayons récolté auprès d'un expert un ensemble de connaissances structurelles sur les classes recherchées dans une image satellite. Comment utiliser ces connaissances afin de mieux classifier l'image ? On est tout d'abord confronté à un problème de représentation; il faut ensuite être capable d'évaluer l'adéquation entre les connaissances et les objets reconnus dans l'image ; enfin, il faut savoir comment exploiter l'aptitude précédente dans le processus de classification. Nous discutons brièvement sur les deux premiers points dans le paragraphe 2, mais la contribution essentielle de l'article se situe au niveau du dernier point. Une approche par modification de régions au moyen de méthodes heuristiques d'optimisation combinatoire est présentée dans le paragraphe 3. L'approche est validée dans le paragraphe 4, où nous présentons quelques résultats expérimentaux. La conclusion est donnée dans le paragraphe 5.

2. Représentation des connaissances

Les connaissances expertes sont plus ou moins certaines et plus ou moins précises. Un système d'inférence floue constitue donc un mode naturel de représentation. Les règles de production sont ordinairement de la forme suivante (bien que les parties prémisses et conclusion soient souvent inversées) [4][8][9] :

Si (classe k), alors (V_1 est A_1) et ... et (V_i est A_i) et ... et (V_N est A_N).

La partie conclusion constitue une description du contexte d'implantation de la classe. Elle se compose de propositions élémentaires « V_i est A_i », où V_i désigne une variable et A_i un sous-ensemble flou. Ces propositions sont unies par des «et » (et éventuellement des «ou ») logiques. En général, les A_i sont déterminés soit *a priori*, en interrogeant les experts, soit *a posteriori*, de manière expérimentale [10][11].

2.1. Connaissances ponctuelles

Lorsque les V_i concernent des pixels (c'est-à-dire lorsque chaque pixel peut être considéré indépendamment des autres), la règle de production représente une connaissance dite *ponctuelle* [9]. On peut avoir par exemple la variable « niveau de gris », ou encore la variable « altitude » (si on dispose d'un modèle numérique de terrain). Le degré d'appartenance à A_i de la valeur de V_i obtenue pour un pixel donné P correspond au degré de vérité de la proposition « V_i est A_i ». Il est noté $\mathbf{m}_{A_i}(P)$. Les opérateurs *min* (opérateur le plus couramment associé au «et » logique) et *max* (pour le «ou ») permettent ensuite de combiner les $\mathbf{m}_{A_i}(P)$ et d'obtenir le degré de vérité de « (V_1 est A_1) et ... et (V_i est A_i) et ... et (V_N est A_N) ». Ce degré, noté $q^{CLk}(P)$, indique dans quelle mesure il est *possible* que le pixel P appartienne à la classe k . Par exemple, pour la connaissance «Si (classe 1), alors (V_1 est A_1) et $\{(V_2$ est A_2) ou (V_3 est $A_3)\}$ » on obtient :

$$q^{CL1}(P) = \min\{\mathbf{m}_{A_1}(P), \max\{\mathbf{m}_{A_2}(P), \mathbf{m}_{A_3}(P)\}\}.$$

2.2. Connaissances structurelles

Considérons maintenant une connaissance experte telle que : « la classe 1 se présente principalement sous la forme de petites régions rondes ». Les pixels ne peuvent plus être traités indépendamment les uns des autres. La connaissance est dite *structurelle*. Elle se compose ici de deux connaissances élémentaires, « forme ronde » et « petite région », qui font intervenir implicitement différentes variables mesurables (comme le ratio d'aspect et l'aire). Nous supposons dans cet article que nous savons rattacher à toute connaissance structurelle élémentaire un ensemble pertinent de variables, et que nous savons mesurer ces dernières, quelle que soit la région considérée dans l'image (nette ou floue). Le lecteur pourra par exemple se reporter à [12]. Pour représenter les connaissances structurelles et mesurer leur adéquation avec les objets mis en évidence dans l'image, il est alors possible d'utiliser un système d'inférence floue multisource [7][8][13]. Ainsi, dans un système hiérarchique de règles floues, une connaissance structurelle CS_k sur la classe k peut s'exprimer de la manière suivante :

Niveau 1 : si (classe k), alors (CSE_1) et ... et (CSE_j) et ... et (CSE_N) .

Niveau 2 : si (CSE_j) , alors $(V_1^j \text{ est } A_1^j)$ et ... et $(V_i^j \text{ est } A_i^j)$ et ... et $(V_{M_j}^j \text{ est } A_{M_j}^j)$.

La partie conclusion du niveau 1 de la règle se compose de connaissances structurelles élémentaires CSE_j unies par des « et » (et éventuellement des « ou ») logiques. Dans le niveau 2, chaque CSE_j est exprimée par l'intermédiaire de variables mesurables V_i^j et de sous-ensembles flous A_i^j . Il est à noter que les V_i^j concernent des régions, non plus des pixels. Soit R une région susceptible d'être attribuée à la classe k . Le degré d'appartenance à A_i^j de la valeur de V_i^j obtenue pour R correspond au degré de vérité $\mathbf{m}_{A_i^j}(R)$ de la proposition « V_i^j est A_i^j ». La combinaison logique des $\mathbf{m}_{A_i^j}(R)$ donne $q^{CSE_j}(R)$, le *degré d'adéquation* de la région R à la connaissance structurelle élémentaire CSE_j . Enfin, au niveau 1, la combinaison logique des $q^{CSE_j}(R)$ produit $q^{CS_k}(R)$, le degré d'adéquation de R à la connaissance CS_k . Par exemple, pour la connaissance CS_2 définie par « Si (classe 2), alors (CSE_1) et (CSE_2) », avec « Si (CSE_1) , alors $(V_1^1 \text{ est } A_1^1)$ ou $(V_2^1 \text{ est } A_2^1)$ » et « Si (CSE_2) , alors $(V_1^2 \text{ est } A_1^2)$ », on obtient :

$$q^{CS_2}(R) = \min\{\max\{\mathbf{m}_{A_1^1}(R), \mathbf{m}_{A_2^1}(R)\}, \mathbf{m}_{A_1^2}(R)\}$$

Notons que Binaghi *et al.* [8] utilisent un système semblable de règles pour exprimer des informations contextuelles (au niveau 1) et évaluer ces informations en exploitant des données multisources (au niveau 2). Ajoutons enfin que nous ne décrirons pas ici comment gérer l'emploi d'adverbes (comme « principalement ») dans la formulation des connaissances.

3. Exploitation des connaissances structurelles

Supposons que nous sachions représenter les connaissances structurelles disponibles et évaluer leur degré d'adéquation avec des régions quelconques de l'image (paragraphe 2). Comment utiliser cette aptitude pour mieux classifier l'image ? C'est le problème de l'intégration des connaissances structurelles dans le processus de classification. Nous proposons une approche par modification de régions.

3.1. Approche par modification des régions (AMR)

Supposons que, par un moyen quelconque, nous ayons obtenu une partition de l'image considérée (partition au sens classique du terme). Nous parlerons de pré-classification. À l'examen de cette seule partition, l'expert humain peut, en un seul coup d'œil, remettre en question le label « classe k » attribué à une région R donnée : tout simplement parce que R ne se conforme manifestement pas aux connaissances structurelles qu'il a sur k . Le label a pu être attribué par erreur (par exemple, R est une région de la classe m , dont les caractéristiques radiométriques sont proches de celles de k).

Ou bien, certains pixels de la frontière de R (des pixels mixtes notamment) ont été mal classés. La partition doit alors être révisée. Certains pixels vont devoir être réassignés. Il est nécessaire pour cela d'examiner à nouveau l'image. L'approche par modification des régions (AMR) imite ce processus. Considérons la figure 1(a). Elle représente le produit d'une pré-classification.

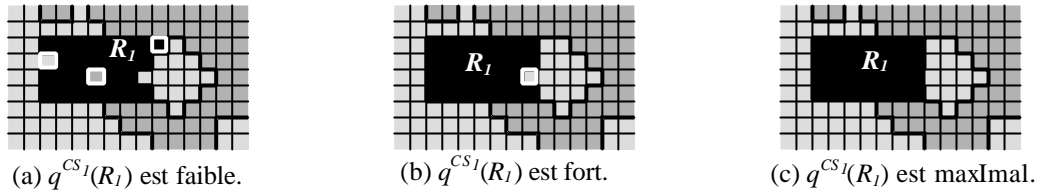


Fig. 1. La région R_1 est-elle rectangulaire ?

Les pixels noirs ont été assignés à la classe 1. Ils constituent une région R_1 . Supposons maintenant que nous disposions de la connaissance CS_1 suivante : « la classe 1 se présente sous la forme de régions rectangulaires ». Il est clair que la réassignation de certains pixels permet d'augmenter significativement $q^{CS_1}(R_1)$, le degré d'adéquation de R_1 à CS_1 (Fig. 1(b)(c)). Bien entendu, si l'on dispose de connaissances structurelles sur les autres classes, cette réassignation peut se traduire dans un même temps par une baisse des degrés d'adéquation des régions voisines. C'est pourquoi nous devons considérer un degré d'adéquation « global » $Q(X)$, où X dénote la partition étudiée. $Q(X)$ est le degré d'adéquation de la partition à l'ensemble des connaissances structurelles disponibles. On peut poser par exemple $Q(X) = \sum_i q_i / n$, où n est le nombre de régions dans la partition et q_i le degré d'adéquation associé à la i -ème région. C'est cette expression qui a été utilisée pour les tests décrits dans le paragraphe 4 (mais d'autres choix sont envisageables, comme $Q(X) = \min_i q_i$). L'objectif de l'AMR est donc de réassigner certains pixels de manière à augmenter globalement le degré d'adéquation des régions. Autrement dit, il s'agit de rechercher une partition X qui maximise $Q(X)$. C'est un problème d'optimisation combinatoire. Pour le résoudre, nous avons considéré trois méthodes heuristiques [14] : les méthodes RS (recuit simulé), MT (méthode tabou), et AG (algorithme génétique), déjà utilisées avec succès dans des problèmes de classification et *clustering* [15][16][17]. Les deux premières méthodes peuvent être décrites très schématiquement par l'algorithme suivant :

-
1. À partir de la solution courante X^{now} , produire une solution candidate alternative X^{cand} .
 2. Si $Q(X^{now}) < Q(X^{cand})$, remplacer X^{now} par X^{cand} .
- Répéter 1. et 2. jusqu'à ce qu'on ne puisse plus trouver de X^{cand} telle que $Q(X^{now}) < Q(X^{cand})$.
-

Bien souvent, la solution finale correspond à un maximum local : elle n'est pas optimale. Pour tenter de résoudre ce problème, des solutions destructives sont parfois acceptées. Il s'agit de solutions X^{cand} qui ne répondent pas à l'inégalité $Q(X^{now}) < Q(X^{cand})$. Dans RS, cette idée se traduit par l'introduction d'un paramètre de température, et le concept de programme de refroidissement. Dans MT, elle se traduit par l'introduction d'une liste tabou, et le critère d'aspiration.

3.2. Production des partitions candidates

En réalité, c'est à une pré-classification floue que nous devons faire appel. La solution initiale X^{init} utilisée par les méthodes heuristiques est une partition nette, mais elle n'est que le produit de la « défuzzification » de la partition floue \tilde{X} issue de la pré-classification. À partir d'une partition X^{now} , solution « du moment », égale au départ à X^{init} , nous devons ensuite produire une partition alternative X^{cand} (dans cet article, sauf indication contraire, le terme « partition » désigne toujours

une partition nette). Pour tout pixel P et toute classe i , notons $\mathbf{m}(P)$ le degré d'appartenance de P à la classe i dans \tilde{X} ; notons $CL^{init}(P)$ la classe de P dans X^{init} et $CL^{now}(P)$ la classe de P dans X^{now} . Pour tout i , nous avons : $\mathbf{m}_{CL^{init}(P)}(P) \geq \mathbf{m}(P)$. La partition alternative X^{cand} est produite à partir de X^{now} en réassignant un pixel P (choisi au hasard) à une classe $CL^{cand}(P) \neq CL^{now}(P)$ (elle aussi choisie au hasard). Il est clair cependant qu'un pixel P satisfaisant $\mathbf{m}_{CL^{init}(P)}(P) \gg \mathbf{m}(P)$ pour tout i différent de $CL^{init}(P)$ ne constituerait pas un choix très judicieux. De même, $CL^{cand}(P)$ ne devrait pas satisfaire $\mathbf{m}_{CL^{init}(P)}(P) \gg \mathbf{m}_{CL^{cand}(P)}(P)$. La partition X^{cand} est donc obtenue de la manière suivante :

[Étape 1] Choisir au hasard un pixel P tel que $\min_{i \neq CL^{init}(P)} \{\mathbf{m}_{CL^{init}(P)}(P) - \mathbf{m}(P)\} \leq \mathbf{s}$ (où \mathbf{s} désigne un seuil prédéfini appartenant à l'intervalle $[0;1]$). **[Étape 2]** Choisir au hasard une classe $CL^{cand}(P)$ telle que $\mathbf{m}_{CL^{init}(P)}(P) - \mathbf{m}_{CL^{cand}(P)}(P) \leq \mathbf{s}$. Si $CL^{cand}(P) = CL^{now}(P)$, revenir à l'étape 1. **[Étape 3]** Réassigner P à la classe $CL^{cand}(P)$. La partition obtenue est X^{cand} .

Mais on peut encore affiner le processus : le «hasard» peut être bridé de manière à ce que (i) la probabilité de choisir P soit d'autant plus grande que la valeur $\min_{i \neq CL^{init}(P)} \{\mathbf{m}_{CL^{init}(P)}(P) - \mathbf{m}(P)\}$ est faible, et (ii) la probabilité de choisir $CL^{cand}(P)$ soit d'autant plus grande que $\mathbf{m}_{CL^{init}(P)}(P) - \mathbf{m}_{CL^{cand}(P)}(P)$ est faible. Par exemple, on peut récrire l'étape 2 ainsi :

[Étape 2] Sélectionner $CL^{cand}(P)$ par la méthode de la roulette : la probabilité pour que P soit assignée à une classe k donnée est fixée à $\mathbf{m}_k(P) / \sum_i \mathbf{m}_i^*(P)$, où $\mathbf{m}_i^*(P)$ égale $\mathbf{m}_i(P)$ si $\mathbf{m}_{CL^{init}(P)}(P) - \mathbf{m}(P) \leq \mathbf{s}$ et vaut 0 sinon. Si $CL^{cand}(P) = CL^{now}(P)$, revenir à l'étape 1.

Remarquons que la pré-classification peut très bien être de type possibiliste [18]. La somme $\sum_i \mathbf{m}_i^*(P)$ ne vaut donc pas nécessairement 1, même si $\mathbf{s} = 1$. Notons également que si la pré-classification est nette, son produit X^{init} peut permettre de générer la partition floue nécessaire au calcul des X^{cand} . Soit P un pixel quelconque et k une classe quelconque. Si P appartient à une région R assignée à la classe k dans X^{init} , on peut donner à $\mathbf{m}_k(P)$ une valeur d'autant plus élevée que P est éloigné de la frontière de R ; sinon, on peut donner à $\mathbf{m}_k(P)$ une valeur d'autant plus faible que P est éloigné de toute région assignée à la classe k . L'AMR va alors avoir tendance à arrondir les régions que l'on s'attend à voir plutôt rondes, à allonger celles que l'on s'attend à voir allongées, etc. On retrouve un mécanisme très semblable dans le domaine de la production de cartes où, souvent, on préfère sacrifier le détail au profit de la lisibilité.

3.3. Modification de régions et recuit simulé

Initialisation : fixer le compteur d'itération ($t=1$), la température initiale T_0 et le programme de refroidissement ($\mathbf{a}(T) = T_0 / t$). Soit X^{init} la partition issue de la phase de pré-classification et soit $X^{now} = X^{init}$. Calculer $Q(X^{now})$.

Itération 1 :

Itération 2 :

1. Produire à partir de X^{now} une partition candidate X^{cand} (voir paragraphe 3.2).
2. Calculer $\mathbf{d} = Q(X^{now}) - Q(X^{cand})$. Si $\mathbf{d} < 0$, poser $X^{now} = X^{cand}$. Sinon, choisir au hasard un élément r de $[0;1]$, et si $r < \exp(-\mathbf{d}/T)$ poser $X^{now} = X^{cand}$.

Répéter **Itération 2** un nombre prédéfini de fois, puis poser $t=t+1$ et mettre à jour la température : $T = \mathbf{a}(T)$.

Répéter **Itération 1** un nombre prédéfini de fois, ou jusqu'à ce que $Q(X^{now})$ atteigne une valeur prédéfinie. X^{now} est la partition finale.

3.4. Modification de régions et méthode tabou

La production à partir de X^{now} d'une partition candidate X^{cand} peut se décrire par une matrice $[P, CL^{now}(P), CL^{cand}(P)]$ (voir paragraphe 3.2). Cette matrice 1×3 est dite *d'information*. Dans l'algorithme de l'AMR par MT, lorsqu'un pixel P est réassigné, la matrice correspondante est rangée dans une *liste tabou*. Il s'agit d'interdire, pendant quelques itérations, la réassignation qui vient d'être effectuée.

Initialisation : fixer le nombre NC de partitions candidates et la taille maximale TM de la liste tabou. Soit $TC=0$ la taille courante de cette liste, X^{init} la partition issue de la phase de pré-classification, et $X^{best} = X^{now} = X^{init}$. Calculer $Q^{best} = Q(X^{now})$.

Itération :

1. Produire NC partitions candidates $X_1^{cand}, \dots, X_{NC}^{cand}$. Soit $\mathbf{m}_1^{cand}, \dots, \mathbf{m}_{NC}^{cand}$ les matrices d'information correspondantes. Calculer $Q(X_1^{cand}), \dots, Q(X_{NC}^{cand})$.
2. Réordonner les $X_1^{cand}, \dots, X_{NC}^{cand}$ et $\mathbf{m}_1^{cand}, \dots, \mathbf{m}_{NC}^{cand}$ de manière à ce que les degrés d'adéquation correspondants $Q(X_1^{cand}), \dots, Q(X_{NC}^{cand})$ apparaissent par valeurs décroissantes. Soit $X_{[1]}^{cand}, \dots, X_{[NC]}^{cand}$, $\mathbf{m}_{[1]}^{cand}, \dots, \mathbf{m}_{[NC]}^{cand}$, $Q_{[1]}^{cand}, \dots, Q_{[NC]}^{cand}$ les éléments ainsi obtenus.
3. Si $\mathbf{m}_{[1]}^{cand}$ n'est pas tabou, ou si $\mathbf{m}_{[1]}^{cand}$ est tabou mais $Q_{[1]}^{cand} > Q^{best}$ (critère d'aspiration), alors poser $X^{now} = X_{[1]}^{cand}$ et $\mathbf{m}^{tabou} = \mathbf{m}_{[1]}^{cand}$, puis aller en 4. Sinon, poser $X^{now} = X_{[b]}^{cand}$ et $\mathbf{m}^{tabou} = \mathbf{m}_{[b]}^{cand}$, où $Q_{[b]}^{cand}$ est la plus grande des valeurs $Q_{[2]}^{cand}, \dots, Q_{[NC]}^{cand}$, avec $\mathbf{m}_{[b]}^{cand}$ non tabou.
4. Ajouter \mathbf{m}^{tabou} en queue de liste et poser $TC = TC + 1$. Si TC dépasse TM , ôter l'élément situé en tête de liste et poser $TC = TM$. Si $Q^{best} < Q(X^{now})$, poser $X^{best} = X^{now}$ et $Q^{best} = Q(X^{now})$.

Répéter **Itération** un nombre prédéfini de fois, ou jusqu'à ce que Q^{best} atteigne une valeur prédéfinie. X^{best} est la partition finale.

3.5. Modification de régions et algorithme génétique

La résolution par algorithme génétique d'un problème d'optimisation passe par la création d'organismes artificiels appelés *individus*, que des opérateurs de reproduction, de croisement, et de mutation permettent de faire évoluer. Nous ne nous étendrons pas ici sur ces opérateurs. Dans l'algorithme de l'AMR par AG, un individu est une partition. La phase d'initialisation consiste à générer une population initiale. Cette population évoluera au cours du déroulement de l'algorithme.

Initialisation : fixer le nombre NC d'individus (partitions candidates), et la probabilité PM de mutation d'un pixel. Produire NC individus $X_1^{cand}, \dots, X_{NC}^{cand}$ (voir paragraphe 3.2).

Itération :

Calculer $Q(X_1^{cand}), \dots, Q(X_{NC}^{cand})$, puis faire évoluer la population : les $X_1^{cand}, \dots, X_{NC}^{cand}$ se reproduisent selon une loi régie par la méthode de la roulette (les probabilités étant définies ici à partir des grandeurs de Q) ; des couples se croisent au hasard ; des mutations interviennent avec la probabilité PM .

Répéter **Itération** un nombre prédéfini de fois, ou jusqu'à ce qu'un individu X^{best} ayant un degré d'adéquation Q^{best} suffisamment élevé soit généré. X^{best} est la partition finale.

Remarquons que les individus considérés sont des individus 2D, et non 1D comme c'est habituellement le cas. Le croisement de deux individus est défini par une ligne, non par un point. Cette ligne est fixée dans l'image de manière aléatoire, soit horizontalement, soit verticalement.

4. Résultats expérimentaux

4.1. Données de synthèse

Nous supposons que la partition floue de la figure 2(a) est le résultat, avant « défuzzification », d'une pré-classification entre 7 classes de 161 éléments ordonnés. La partition nette X^{init} associée à cette partition floue est représentée par la figure 2(b). La figure 2(c) montre le résultat de la classification manuelle effectuée par l'expert humain (en fait, pour révéler le dessous des cartes, nous avons construit toutes les données de synthèse à partir de cette partition « experte »). Nous considérons d'autre part les connaissances structurelles suivantes : les classes 1, 2 et 7 se présentent sous la forme de petites régions ; les classes 4 et 6 sous la forme de régions de taille moyenne ; les classes 3 et 5 sous la forme de régions très étendues ; enfin, pour tout $k \neq 1$, une région de la classe k est bordée du côté gauche par une région de la classe $(k-1)$. La construction de ces données et connaissances de synthèse s'inspire du problème de la classification géomorphologique d'un atoll. Les données représentent les pixels d'une coupe transversale de la couronne récifale de l'atoll [19]. Le long de la coupe, de l'océan vers le lagon, les classes géomorphologiques ne sont pas situées dans n'importe quel ordre. On trouve communément la classe « crête externe », puis « dalle du platier », puis « conglomérat récifal », etc., puis « platier interne ». Pour représenter ces connaissances, adoptons un système hiérarchique de règles floues, comme dans le paragraphe 2.2. À chaque classe est associée une règle à deux niveaux. Le niveau 1 fait intervenir deux connaissances structurelles élémentaires : l'une concerne la taille des régions, l'autre les relations spatiales entre régions.

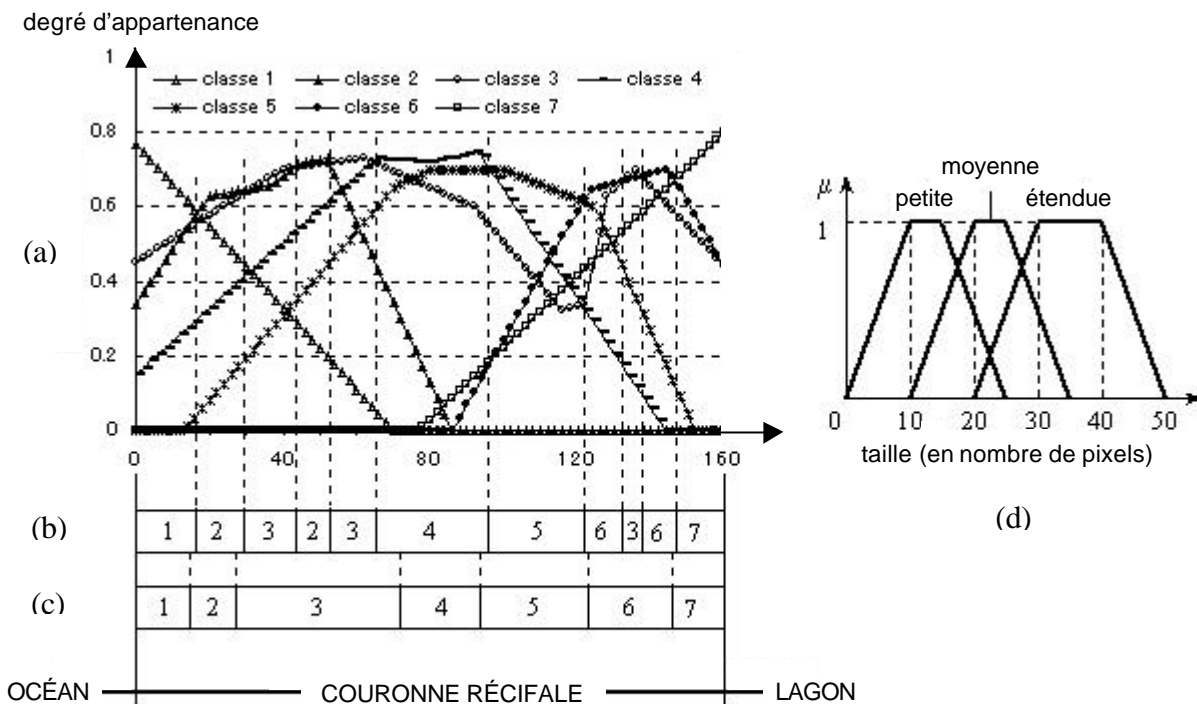


Fig. 2. Données de synthèse. (a) Partition floue issue de la phase de pré-classification. (b) Partition nette correspondante. (c) Partition nette issue d'une classification manuelle. (d) Valeurs linguistiques de la variable « taille ».

Nous avons ainsi : « Si (classe 2), alors «petite région » et « la région voisine de gauche est de classe 1 » ». Au niveau 2, on peut rattacher à la première connaissance élémentaire CSE_1 la variable mesurable « nombre de pixels ». Pour calculer le degré d'adéquation $q^{CSE_1}(R)$ de CSE_1 avec une région R donnée, on peut ensuite utiliser la variable linguistique « taille » définie par la figure 2(d). Le degré d'adéquation $q^{CSE_2}(R)$ de CSE_2 avec R est encore plus facile à évaluer. Il est fixé à 1 si R est bordée sur sa gauche par une région de la classe $k-1$ (où k désigne la classe à laquelle R a été assignée), et à 0 sinon. Finalement, la règle floue associée à la classe 2 permet d'attacher à R le degré d'adéquation $\min\{q^{CSE_1}(R), q^{CSE_2}(R)\}$. On voit à ce stade que pour toute partition X , il est possible d'évaluer $Q(X)$, le degré d'adéquation de X à l'ensemble des connaissances structurelles disponibles (cf. paragraphe 3.1). Pour X^{init} , on obtient la valeur 0,32. Pour la partition censée représenter la réalité (Fig. 2(c)), c'est 0,96. Le processus de modification des régions par recuit simulé a été exécuté 10 fois, les processus de modification par méthode tabou et algorithme génétique également. La figure 3 montre les résultats obtenus. À chaque itération, le degré d'adéquation de la partition s'élève. La valeur maximale atteinte par Q est 0,96. Cette valeur est atteinte dans tous les cas, c'est-à-dire 10 fois sur 10, quelle que soit la méthode heuristique utilisée. Néanmoins, beaucoup plus d'itérations sont nécessaires pour l'AMR par AG. Notons que les partitions finales ne correspondent pas exactement à la partition attendue (celle de la figure 2(c)). Mais le pourcentage des pixels bien classés est toujours supérieur à 82,6%, qui est la valeur obtenue pour la partition initiale (celle de la figure 2(b)). En moyenne, il est supérieur à 99% pour l'AMR par RS ou MT, et supérieur à 96% pour l'AMR par AG. En conclusion, c'est en utilisant la méthode du recuit simulé que l'approche par modification de régions exploite le plus efficacement les connaissances structurelles disponibles.

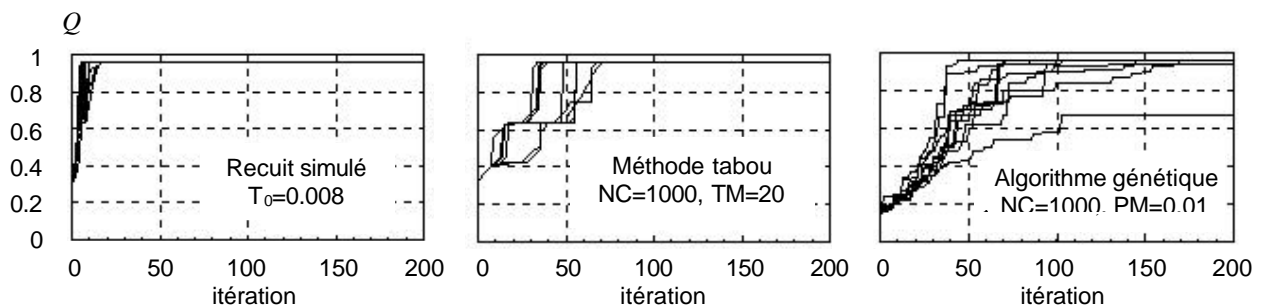


Fig. 3. Résultats des expérimentations sur les données de synthèse.

4.2. Données réelles

Nous préparons actuellement une série de tests sur des images satellites SPOT XS d'atolls. Pour classifier ce type d'images, l'expert photo-interprète fait intervenir implicitement de nombreuses connaissances structurelles (certaines sont d'ailleurs relativement complexes). Dans cet article, et pour nos premiers tests sur données réelles, nous considérons une image très simple, dans laquelle peu de classes sont présentes, et dont l'analyse ne requiert pas l'aide d'experts en géomorphologie récifale. Il s'agit de l'image RVB d'un couteau avec trois rivets, comme le montre la figure 4(a). Quatre classes sont recherchées : « fond », « ombre », « manche » et « rivet ». Une première classification de cette image (classification floue supervisée que nous ne décrivons pas ici) a permis d'obtenir la partition représentée par la figure 4(b). Un examen minutieux de l'image originale montre que 29% des pixels « rivet » sont mal classés. La figure étant petite, il est difficile de bien s'en rendre compte. Malgré tout, on perçoit nettement des anomalies sur le rivet de gauche. Dans le but d'améliorer ce premier résultat, la connaissance structurelle suivante est introduite : « un rivet est une petite région ronde ». Pour exprimer la connaissance élémentaire « région ronde » nous

faisons appel aux variables «ratio d'aspect » et «densité » (Fig. 4(d)(e)). Pour « petite région », nous faisons naturellement intervenir la variable «aire » (Fig.4(f)). La méthode du recuit simulé est la seule à être utilisée ici. La figure 4(c) représente la partition obtenue après 2500 itérations. Le pourcentage des pixels «rivet » mal classés est tombé à 19%. Le degré d'adéquation de la partition atteint 0,95, contre 0,61 pour la partition initiale (Fig. 4(b)). Aussi élevé que paraisse ce degré, les rivets obtenus ne sont pas parfaitement circulaires. Cela est dû en partie à la manière dont la connaissance « région ronde » a été représentée. Il est clair que la qualité des résultats produits par l'approche de modification des régions dépend non seulement de la qualité des partitions floue et nette initiales, mais aussi de la qualité des connaissances, et de la représentation qui en a été faite.

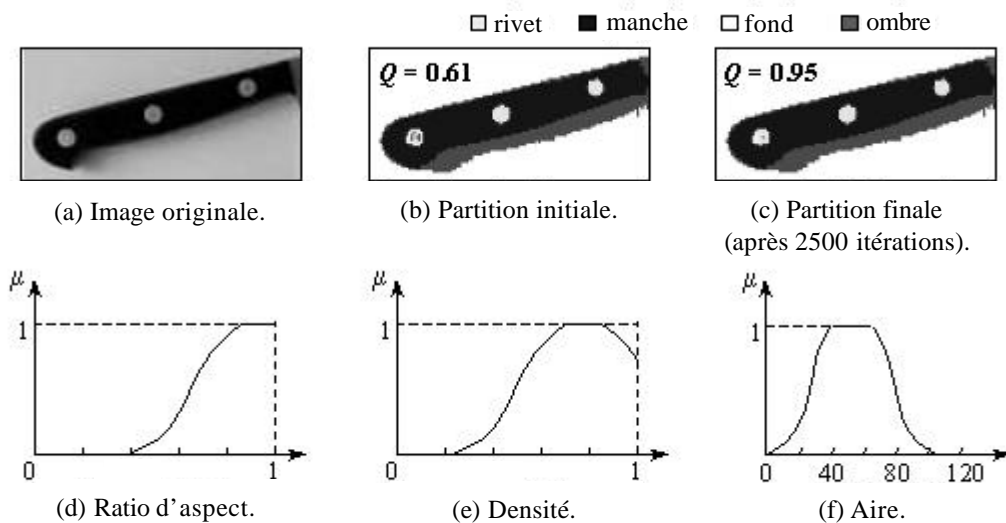


Fig. 4. Données réelles.

5. Conclusion

Dans cet article, nous avons décrit une approche pour améliorer la classification automatique d'images satellites en exploitant des connaissances structurelles d'experts. Elle est fondée sur : (i) la représentation des connaissances par un système d'inférence floue ; (ii) le calcul, dans une étape préliminaire, d'une partition floue et de la partition nette qui lui est associée ; (iii) l'utilisation de méthodes heuristiques d'optimisation combinatoire. Nous avons présenté les résultats de premiers tests menés sur des données et connaissances de synthèse, ainsi que sur l'image RVB d'un objet simple très structuré. Les partitions nettes issues de la pré-classification sont modifiées de manière cohérente, par réassignation correcte de pixels initialement mal classés. Ces résultats encourageants valident notre approche. Nous allons maintenant effectuer un premier ensemble d'optimisations algorithmiques. Dans le cas de l'image RVB par exemple, qui est de très petite taille (115×54), 35% des pixels «rivet » initialement mal classés ont été récupérés. Mais ce résultat a nécessité près d'une heure de calcul, une performance très médiocre qu'il faut néanmoins relativiser. Tout d'abord, les programmes ont été écrits en Java et exécutés sur un MMX Pentium 300MHz. Ensuite, dans les algorithmes actuels, les modifications successives apportées aux partitions sont minimales, puisqu'un seul pixel distingue une partition candidate de la partition dont elle est issue. Surtout, chaque partition candidate est entièrement réétiquetée et tous les degrés d'adéquation sont recalculés, alors que la grande majorité des régions restent manifestement inchangées. Au-delà des questions d'optimisation, nous préparons une série de tests sur des images satellites d'atolls. Nous envisageons également d'appliquer notre approche au problème de la détection de cibles.

Remerciements

Nous tenons à remercier Didier Dubois et Henri Prade pour leurs précieux conseils et commentaires.

Références

- [1] J. A. Richards, D. A. Landgrebe et P. H. Swain, « A means for utilizing ancillary information in multispectral classification », *Remote Sensing Environ.*, no. 12, pp. 463-477, 1982.
- [2] C. M. Gurney et J. R. G. Townsend, « The use of contextual information in the classification of remotely sensed data », *Photogram. Eng. Remote Sensing*, no. 49, pp. 55-64, 1983.
- [3] L. A. Zadeh, « Fuzzy sets », *Information and Control*, no. 8, pp. 338-353, 1965.
- [4] P. Clark, C. Feng, S. Matwin et K. Fung, « Improving image classification by combining statistical, case-based and model-based prediction methods », *Fundamenta Informatica*, vol. 30, no. 3-4, pp. 227-240, 1996.
- [5] E. H. Zahzah et J. Desachy, « Numeric and symbolic data combination for satellite image interpretation », *IGARRS'93*, Tokyo, Japon, 1993, pp. 1704-1706.
- [6] D. P. Mandal, C. A. Murthy et S. K. Pal, « Analysis of IRS imagery for detecting man-made objects with a multivalued recognition system », *IEEE Trans. Sys. Man. Cyber.*, vol. 26, no. 2, pp. 241-247, 1996.
- [7] B. Solaiman, L. E. Leland et F. T. Ulaby, « Multisensor data fusion using fuzzy concepts: Application to land-cover classification using ERS-1/JERS-1 SAR composites », *IEEE Trans. Geosci. Remote Sensing*, vol. 37, no. 3, pp. 1316-1326, 1999.
- [8] E. Binaghi, P. Madella, M.G. Montesano et A. Rampini, « Fuzzy contextual classification of multisource remote sensing images », *IEEE Trans. Geosci. Remote Sensing*, vol. 35, no. 2, pp. 326-339, 1997.
- [9] J. Desachy, *Interprétation automatique d'images satellite: le système ICARE*, Thèse de Doctorat d'Etat de l'Université Paul Sabatier, Toulouse III, 1991.
- [10] J. L. Chameau et J. C. Santamarina, « Membership functions Part I: Comparing methods of measurement », *Int. J. Approximate Reasoning*, no. 1, pp. 287-301, 1987.
- [11] I. B. Türksen, « Measurement of membership functions and their acquisition », *Fuzzy Sets and Systems*, vol. 40, no. 1, pp. 5-38, 1991.
- [12] A. Rosenfeld, « Fuzzy geometry: an updated overview », *Information Sciences*, vol. 110, no. 3-4, pp. 127-33, 1998.
- [13] F. Russo et G. Ramponi, « Fuzzy methods for multisensor data fusion », *IEEE Trans. Instrumentation and Measurement*, vol. 43, no. 2, pp. 288-294, 1994.
- [14] E. H. L. Aarts et J. K. Lenstra, *Local Search in Combinatorial Optimization*, Wiley Interscience Series in Discrete Mathematics and Optimization, John Wiley & Son Ltd., juin 1997.
- [15] M. Bao, « Classification of multi-temporal SAR images and INSAR coherence images using adaptive neighborhood model and simulated annealing approach », *20th. Asian Conference for Remote Sensing*, Hong Kong, Chine, 1999.
- [16] K. S. Al-Sultan, « A tabu search approach to the clustering problem », *Pattern Recognition*, vol. 28, no. 9, pp. 1443-1451, 1995.
- [17] M. Sonka, S. K. Tadikonda et S. M. Collins, « Knowledge-based interpretation of MR brain images », *IEEE Trans. Med. Imag.*, vol. 15, no. 4, pp. 443-452, 1996.
- [18] R. Krishnapuram et J. M. Keller, « A possibilistic approach to clustering », *IEEE Trans. Fuzzy Systems*, vol. 1, no. 2, pp. 98-110, 1993.
- [19] S. Andréfouët, L. Roux, Y. Chancerelle et A. Bonneville, « A fuzzy-possibilistic scheme of study for objects with indeterminate boundaries: application to French Polynesian reefs », *IEEE Trans. Geosci. Remote Sensing*, vol. 38, no. 1, pp. 257-270, 2000.