

Classification d'image satellitaire hyperspectrale en zone rurale et périurbaine

Olivier PONY^{1&2}, Umberto POLVERINI¹, Laurent GAUTRET¹, Josiane ZERUBIA², Xavier DESCOMBES²

¹ASTRIUM, département Modélisation, Outils, Simulation

31, rue des Cosmonautes

Z.I. du Palays

31402 Toulouse Cedex 4

²INRIA, projet Ariana, commun CNRS/INRIA/UNSA

2004, route des Lucioles

BP 93

06902 Sophia Antipolis

olivier.pony@astrium-space.com

Résumé – Les projets futurs d'observation satellitaire hyperspectrale à des résolutions d'une dizaine de mètres fourniront des images de zones rurales et périurbaines. L'information spectrale précise du terrain permet une classification d'un tel paysage. Afin de régulariser une classification par maximum de vraisemblance, nous considérons une modélisation de l'image par champs de Markov. Nous utilisons une expression de son énergie selon le modèle de Potts, que nous avons rendu adaptatif aux classes, pour compenser la sensibilité des structures fines (routes) à une régularisation. De plus, l'étude de la texture fournit des critères supplémentaires caractéristiques de certaines classes (forêt, paysage urbain). Une image et une vérité terrain ont permis de mettre en évidence les méthodes, les critères et le paramétrage adéquats pour optimiser une classification.

Abstract – The future projects in hyperspectral satellite Earth observation with a resolution of ten meters will provide images from rural and semiurban regions. Precise spectral information allows a classification of such a landscape. In order to smoothen a maximum likelihood classification, we consider a Markov random field image modelling. As an expression of its energy, we use the Potts model, which we made adaptive to classes, to compensate the sensitivity of some thin structures (roads) to a regularisation. Moreover, texture supplies new criteria, which characterise some classes (forests, urban landscape). One image and a ground truth enabled us to highlight the best methods, criteria, and parameters, to optimise a classification.

1. Introduction

1.1 L'imagerie satellitaire

Les satellites d'observation fournissent et fourniront de plus en plus des images de la Terre à des résolutions spatiales et spectrales adéquates pour un paysage rural ou périurbain. Une exploitation évidente de telles données est la classification automatique dans le cadre d'un plan d'occupation des sols [6]. Sur des images satellitaires simulées, nous avons évalué et enrichi les méthodes classiques de classification, en s'intéressant en particulier à des images à 10 canaux spectraux et 10 mètres de résolution.

1.2 Classification d'une scène agricole

Dans cet article, nous présentons une image satellitaire simulée de zone agricole et forestière proche de Boigneville dans l'Essonne. Elle comporte 10 bandes spectrales allant du bleu (440 nm) au proche infra-rouge (820 nm). Sa résolution est de 10 m. On y trouve diverses cultures : blé, maïs, orge, colza, pois, betteraves, ainsi que d'autres types de paysages : jachère, forêt, route.

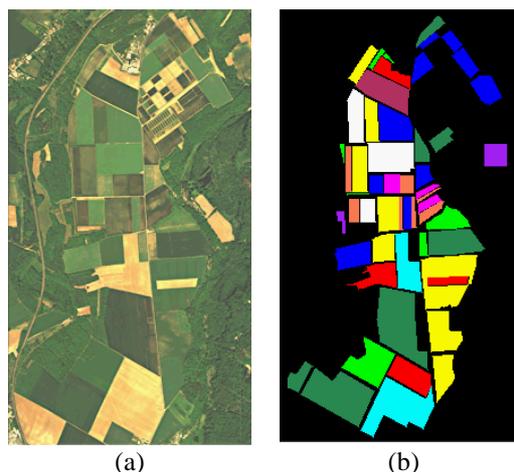


FIG. 1 : (a) Image originale et (b) vérité terrain

Nous disposons sur cette image d'une vérité terrain qui nous permettra de valider et d'évaluer notre classification.

Nous avons appliqué une classification par maximum de vraisemblance sur cette scène, dont le résultat est présenté à la figure 4a. L'hypothèse d'indépendance d'un pixel vis-à-vis de son environnement est source de confusion dans la

classification de certains pixels. Afin de régulariser notre classification, nous allons appliquer un recuit simulé sur l'image modélisée par champs de Markov [3] avec le modèle d'énergie de Potts. Le taux d'erreur est ainsi divisé par 2.

Enfin, l'ajout de critères de texture adaptés, doublé d'une modification du modèle de Potts qui le rend adaptatif aux classes, nous permet d'améliorer le taux d'erreur de la première classification d'un facteur 5 [9].

2. Classification

Le but de toute classification est d'attribuer à chaque pixel d'une image une classe choisie parmi une collection de classes (ex : blé, soja, maïs, etc.). Ici, nous appliquerons une classification par maximum de vraisemblance, supervisée, c'est-à-dire que le nombre de classes et leurs caractéristiques sont fixés au départ.

2.1 Définitions, notations

Une image S est constituée de sites s_i ou pixels. Une classe est notée λ_i , l'ensemble des classes est Λ . La classe attribuée au pixel s est notée x_s . L'ensemble des valeurs que prend un pixel s sur les critères sélectionnés (radiométrique, texture) est un état y_s . L'espace des états E est l'ensemble de tous les états possibles. Sa dimension est le nombre de critères.

L'état global d'une image est noté y , et sa classification globale est notée x . Soient X et Y les champs aléatoires associés à x et y .

2.2 Maximum de vraisemblance

En un pixel s , on désire pouvoir exprimer en fonction de y_s la probabilité d'appartenance du pixel à chacune des classes λ_i . On fait l'hypothèse que la loi de probabilité $P(y_s/x_s = \lambda_i)$ est gaussienne. Il suffit donc pour chaque classe d'estimer un vecteur moyenne μ_i et une matrice de covariance Σ_i dans l'espace des états E . C'est la phase d'apprentissage des classes où l'on calcule ces statistiques sur des échantillons représentatifs.

On fait également une hypothèse d'indépendance conditionnelle des probabilités, c'est-à-dire que l'état d'un pixel, connaissant sa classe, est indépendant des états des autres pixels, soit :

$$P(Y=y/X=x) = \prod_s P(Y_s=y_s/X_s=x_s) \quad (1)$$

Le problème global de classification optimale s'écrit :

$$x_{opt} = \underset{x}{\operatorname{argmax}} P(X=x/Y=y) \quad (2)$$

Les deux hypothèses précédentes, et celle que toutes les classifications x sont *a priori* équiprobables, nous ramènent à un problème local simplifié :

$$x_{s \text{ opt}} = \underset{\lambda_i}{\operatorname{argmin}} \left((y_s - \mu_i)^T [\Sigma_i]^{-1} (y_s - \mu_i) + \log \det(\Sigma_i) \right) \quad (3)$$

2.3 Champs de Markov - Modèle proposé

2.3.1 Topologie

On définit sur l'image S un système de voisinage faisant correspondre à chaque site s un ensemble de pixels voisins V_s . La figure 2 présente deux exemples naturels de systèmes de voisinage.



FIG. 2 : Systèmes de voisinage en 4 et 8 connexité

La figure 3 illustre un système de voisinage plus élaboré, que l'on nomme topologie à trous [2]. Nous verrons plus tard son intérêt.

Une clique est un ensemble de pixels voisins 2 à 2. Le nombre de pixels ou sites d'une clique est 1 ou 2 en 4-connectivité, et 1, 2, 3 ou 4 en 8-connectivité.

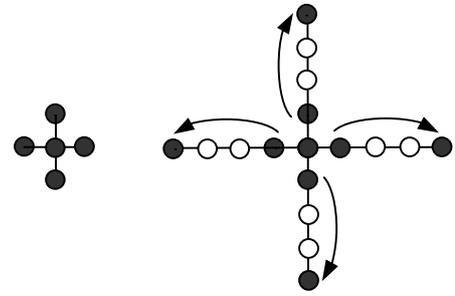


FIG. 3 : Systèmes de voisinage classique et à trous

2.3.2 Champs de Markov et modèle de Potts

Notons x^s la configuration de l'image exceptée le site s . Un champ de Markov X sur l'image S relativement à un système de voisinage V vérifie :

$$\begin{aligned} \forall x, P(X=x) > 0 \\ \forall s, P(X_s=x_s/x^s) = P(X_s=x_s/V_s) \end{aligned} \quad (4)$$

La probabilité pour un pixel d'appartenir à une classe donnée ne dépend que de l'état des pixels voisins. Les champs de Markov sont, d'après le théorème de Hammersley-Clifford, équivalents aux champs de Gibbs. Les champs de Gibbs font intervenir une énergie U_c sur les cliques c :

$$P(X=x) = Z^{-1} \exp\left(-\sum_c U_c(x)\right) \quad (5)$$

Ces deux formulations équivalentes conduisent à la suivante :

$$P(X_s=x_s/X^s=x^s) = \frac{\exp\left(-U_s(x_s/V_s)\right)}{\sum_{x_s \in \Lambda} \exp\left(-U_s(x_s/V_s)\right)} \quad (6)$$

Cette dernière formule nous permet de simuler aisément des champs de Markov, après avoir adopté un modèle d'énergie. Le modèle que nous utilisons ici est le modèle de Potts. Dans ce modèle, seules les cliques d'ordre 1 (les sites)

et d'ordre 2 (les couples de voisins) ont une énergie non nulle.

Pour les cliques d'ordre 1, on reprend le terme d'attache aux données de la formule (3). Pour un site s et une classe λ :

$$U_s(\lambda) = (y_s - \mu_\lambda)^T [\Sigma_\lambda]^{-1} (y_s - \mu_\lambda) + \log \det(\Sigma_\lambda) \quad (7)$$

Pour les cliques d'ordre 2, on introduit un terme de régularisation, avec $\beta \geq 0$:

$$U_{c=(s,t)}(x_s, x_t) = \begin{cases} -\beta & \text{si } x_s = x_t \\ +\beta & \text{si } x_s \neq x_t \end{cases} \quad (8)$$

L'énergie globale d'une classification est donc :

$$U(x/y) = \sum_s \left((y_s - \mu_{x_s})^T [\Sigma_{x_s}]^{-1} (y_s - \mu_{x_s}) + \log \det(\Sigma_{x_s}) \right) + \beta \sum_{(s,t) \in C_2} \delta_{x_s, x_t} \quad (9)$$

où δ vaut -1 si $x_s = x_t$, +1 sinon. β est donc un paramètre de pondération du critère de régularisation.

2.3.3 Recherche de l'optimum par ICM ou recuit simulé

L'objectif est de trouver la configuration optimale, celle qui maximise la probabilité, donc qui minimise l'énergie. L'ICM (*Iterated Conditional Mode*) est une méthode simple qui donne une classification sous-optimale [1]. Partant d'une configuration initiale, elle consiste à balayer les pixels un à un, et d'y choisir la classe qui minimise localement l'énergie. Elle est rapide, mais dépend fortement des conditions initiales et converge vers le premier minimum local rencontré.

Nous utiliserons de préférence une méthode plus lente, dérivée de l'algorithme de Métropolis [8], qui converge vers le minimum d'énergie global : le recuit simulé. À chaque itération, on choisit en chaque pixel une classe au hasard, qui lui est affectée si cela réduit l'énergie localement. Si la classe choisie augmente l'énergie, alors elle est acceptée avec une probabilité d'autant plus forte que ce saut d'énergie est faible. Cette tolérance est par ailleurs réglée par un terme de température qui décroît lentement au cours des itérations. Une température forte induit une tolérance plus forte à une augmentation de l'énergie.

3. Application

Nous avons appliqué la théorie précédente pour classifier l'image de Boigneville. Cela nous permet de déterminer les méthodes les plus performantes.

3.1 Critères d'évaluation

Disposant d'une vérité terrain, il nous est possible de calculer, après une classification, une matrice de confusion des classes entre elles. On peut également visualiser les zones à problème. Pour comparer simplement nos différentes approches, nous nous contenterons d'un simple critère de pourcentage de pixels mal classifiés.

3.2 Régularisation

Pour l'image de Boigneville, il ressort que la classification est de meilleure qualité pour une topologie de type voisinage

en 8-connexité à trous. La topologie à trous permet de détruire les petits ensembles de pixels mal classifiés, dont la proximité des uns aux autres les rend résistants à la régularisation.

La qualité de la classification dépend également de la pondération entre le terme d'attache aux données et le terme de régularisation, dans l'expression de l'énergie, formule (9). Le paramètre β fixe cette pondération. Cela correspond à un paramètre géométrique. En effet, plus β est élevé, plus on donne d'importance à la cohérence entre pixels voisins. On obtient alors des vastes parcelles homogènes. Les structures de petites tailles sont éliminées. Inversement, un β trop faible bride la régularisation, voir figure 4.

Dans l'image de Boigneville, la valeur optimale du paramètre β vaut environ 2. Cet optimum est un compromis entre la taille des petites parcelles que l'on souhaite préserver et celle des objets que l'on souhaite éliminer.

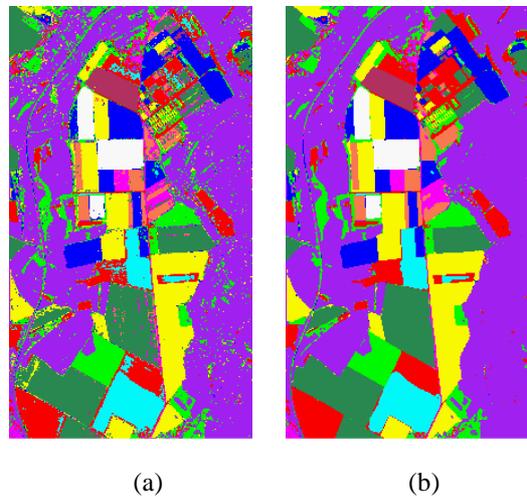


FIG. 4 : (a) Classification brute, $\beta=0$ et (b) régularisation avec une topologie en 8-connexité à trous, $\beta=3$

4. Régularisation adaptative

Le modèle appliqué précédemment présente des faiblesses pour restituer certaines classes. Dans notre image, à une résolution de 10 m, les routes ont une épaisseur de 1 ou 2 pixels. Un pixel de la route traversant la forêt à gauche de l'image possède davantage de voisins de la classe forêt que de la classe route. Ainsi, le terme de régularisation peut inciter un tel pixel à être reclassé dans la classe forêt. Nous avons fait évoluer le modèle d'énergie de Potts pour tenir compte de la sensibilité de ces classes.

4.1 Modèle de Potts modifié

Dans la formule (9), le terme de régularisation est modifié. En plus d'un coefficient pondérateur global β , on introduit un coefficient β_λ pour chacune des classes. β_λ est d'autant plus fort que la classe est sensible à la régularisation.

$$U(x/y) = \sum_s \left((y_s - \mu_{x_s})^T [\Sigma_{x_s}]^{-1} (y_s - \mu_{x_s}) + \log \det(\Sigma_{x_s}) \right) + \beta \sum_{(s,t) \in C_2} \beta_{x_s} \beta_{x_t} \delta_{x_s, x_t} \quad (10)$$

4.2 Application

Dans notre cas, on fixe $\beta_\lambda = 1$ pour toutes les classes, et $\beta_\lambda = 4$ pour la classe route. Ainsi, dans un voisinage, la présence d'un ou deux pixels de route est un fort critère pour l'affectation du pixel courant à cette classe route.

Nous avons testé, pour un système de voisinage en 8-connexité classique (sans trous), l'effet de la régularisation sur la route qui traverse la forêt, d'une part avec le modèle de Potts classique, et d'autre part avec le modèle de Potts adaptatif aux classes, voir figure 5.

En outre, la matrice de confusion met en évidence la tendance de certaines classes à être systématiquement associées à une autre classe. L'évolution du modèle d'énergie nous a ainsi permis de corriger une ambiguïté entre les classes « blé d'hiver » et « blé ».

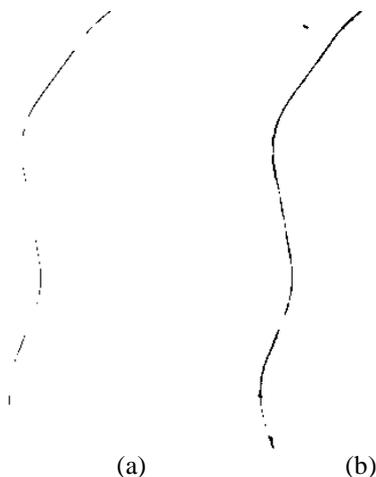


FIG. 5 : Robustesse des routes à (a) une régularisation classique et (b) une régularisation adaptative

5. Critère de texture

L'information contextuelle est exploitée au niveau de la régularisation, grâce au formalisme markovien. On se propose ici d'ajouter à la radiométrie des critères de texture calculés sur les images originales. Ceci présente un intérêt notable pour la classification des paysages urbains ou des forêts.

5.1 Variance locale

Il existe un certain nombre de modèles de texture. Les paramètres extraits des matrices de cooccurrence [4] ou les réponses aux filtres de Gabor [7] sont des critères anisotropes. Ici, nous avons utilisé un critère isotrope inspiré des modèles fractals [10].

En chaque pixel d'une des bandes spectrales de l'image originale, on calcule la variation totale de radiométrie V dans un voisinage figé, à diverses résolutions ou échelles e . On déduit un paramètre d , proche de la notion de dimension de Hausdorff :

$$d-2 = \frac{\partial \log(V_e)}{\partial \log(e)} \quad (11)$$

Une zone homogène donne $d-2 = 0$, alors qu'une surface texturée ou bruitée donne $0 < d-2 < 1$. L'avantage de cette méthode est qu'un bord de champ ou une route se comportent comme une surface homogène, voir figure 6. C'est une méthode qui, appliquée à une bande du proche infrarouge, permet de séparer la forêt et les paysages urbains de la campagne. Dans le spectre visible, ce critère de texture est fort uniquement pour le paysage urbain. Ainsi, en utilisant plusieurs longueurs d'onde, on arrive à séparer trois grands types de paysages : ville, forêt, campagne.

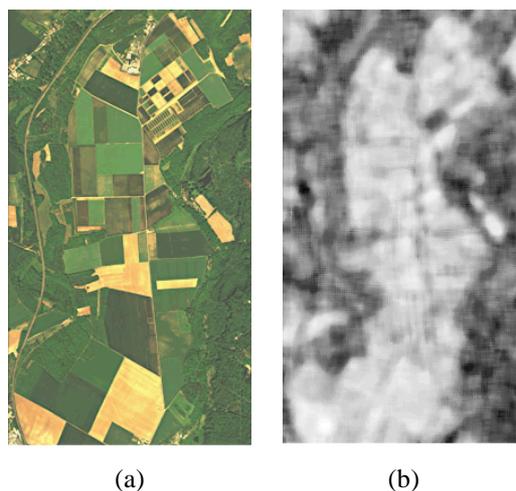


FIG. 6 : (a) Image originale et (b) réponse du critère de texture

5.2 Résultats

Nous avons appliqué une classification en ajoutant aux critères initiaux des critères artificiels de texture, sans régularisation d'abord. Nous avons ensuite effectué la régularisation, en comparant le modèle de Potts original et le modèle adaptatif aux classes.

TAB. 1 : Performances comparées de différentes opérations de régularisation

	Erreur (%)
Classification 10 bandes radiométriques	7,64
Classification + régularisation : opérateurs morphologiques (<i>sieve & clump</i>)	6,49
Classification + régularisation : ICM	4,32
Classification + régularisation : recuit simulé	4,00
Classification 10 bandes + 3 bandes de texture (critère : variance locale)	6,27
Classification 10 bandes + 3 bandes de texture (critère : filtres de Gabor)	6,49
(*) Classification 10 bandes + 2x3 bandes de texture (critères : variance locale + Gabor)	4,77
Classification (*) + régularisation : recuit simulé	2,75
Classification (*) + régularisation : recuit simulé + modèle de Potts adaptatif	1,42

Une simple classification sur l'image initiale nous donne un taux d'erreur de 7,64%. Une régularisation par champs de Markov, selon le modèle de Potts, opérée par un recuit simulé correctement paramétré réduit la confusion à 4,00%.

L'introduction de critères de texture permet déjà d'atteindre, après une simple classification, un taux d'erreur du même ordre, légèrement supérieur. La régularisation précédente apporte ensuite un gain notable (2,75%). Enfin, l'emploi du modèle de Potts adaptatif conduit à un taux de pixels mal classés de 1,42%.

6. Conclusion

Les résultats des traitements successifs appliqués à l'image témoignent de la possibilité d'optimiser certaines méthodes de classification. Cependant, des efforts d'automatisation doivent accompagner la complexification du modèle. Le modèle de Potts adaptatif introduit autant de nouveaux paramètres que de classes définies. Le paramétrage optimal du modèle est fortement lié au contenu de l'image et à l'application.

L'expérience montre que pour des résolutions de 10 m, on parvient à séparer le paysage urbain de la campagne et des forêts. D'une part, le traitement des régions texturées dépend fortement de la résolution de travail. Aussi faut-il envisager l'exploitation de modèles de texture multi-résolution [5]. D'autre part, la restriction de notre cadre de travail exclut de « rentrer dans la ville ».

Enfin, nous nous sommes peu attardés sur la première étape : l'apprentissage. Le choix des classes et des régions d'apprentissage, la hiérarchisation en classes et sous-classes, sont des points d'étude cruciaux pour améliorer la classification initiale.

Références

- [1] Besag (J.). – Spatial interaction and the statistical analysis of lattice systems. *J. Royal Statist. Soc. B*, n36, 1974, pp. 192-326.
- [2] Descombes (X.), Moctezuma (M.), Maître (H.) et Rudant (J.). – Coastline detection by a markovian segmentation on sar images. *Signal Processing*, vol. 55, n1, 1996, pp. 123-132.
- [3] Geman (S.) et Graffigne (C.). – Markov random field image models and their applications to computer vision. *Proceedings of the International Congress of Mathematicians : Berkeley, Août 1986*, pp.1496-1517.
- [4] Haralick (R.). – Statistical and structural approaches to texture. *Proc. of the IEEE*, vol. 67, n5 Mai 1979, pp. 786-804.
- [5] Krishnamachari (S.) et Chellappa (R.). – Multiresolution Gauss-Markov random field models for texture segmentation. *IEEE Trans. on Image Processing*, vol. 6, n2, Février 1997, pp. 251-267.
- [6] Landgrebe (D.). – *Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data*. – Rapport technique, Purdue University, Indiana, 1998.
- [7] Lorette (A.). – *Analyse de texture par méthodes markoviennes et par morphologie mathématique : application à l'analyse des zones urbaines sur des images satellitaires*. – Thèse de doctorat – UNSA, 1999.
- [8] Metropolis (N.), Rosenbluth (A.), Rosenbluth (N.), Teller (A.) et Teller (E.). – Equation of state calculations by fast computing machines. *J. Chemical Physics*, n21, 1953, pp. 1087-1091.
- [9] Pony (O.), Zerubia (J.), Descombes (X.), Classification d'images satellitaires hyperspectrales en zone rurale et périurbaine, *Rapport de recherche – INRIA*, n4008, Sept. 2000.
- [10] Serafim (A.L.). – Segmentation of natural microtextures by joining local and global fractal model parameters. *Proc. SPIE – Int. Soc. Opt. Eng. (USA), Proceedings of the SPIE – The International Society for Optical Engineering*, vol. 3101, 1997, pp. 38-48.