

Stratégie Contextuelle et Extraction de Primitives pour la Segmentation des Images Multispectrales : Etude d'une Simulation

Yuttapong RUNGSUNSERI, Wojciech PIECZYNSKI, Christian ROUX

Groupe Traitement d'Images, Département Mathématiques et Systèmes de Communication
Ecole Nationale Supérieure des Télécommunications de Bretagne, B.P. 832, 29285 BREST Cédex

RESUME

Cette étude traite de l'utilisation conjointe de l'information spectrale et de l'information spatiale pour la segmentation des images multispectrales de satellites. En utilisant une simulation unidimensionnelle, type balayage vidéo, et en prenant un modèle Markovien, on présente des résultats permettant de choisir entre deux types de stratégies de classification : une classification utilisant un contexte spatial et une information spectrale réduites par extraction de primitives, d'une part, et d'autre part, une classification non-contextuelle utilisant l'information spectrale complète. La classification s'opère selon un critère de maximum de vraisemblance dans tous les cas envisagés, l'extraction de primitives étant, le cas échéant, réalisée par l'approche discriminante de Fisher. Les résultats de simulation montrent quelle stratégie choisir en fonction de la corrélation spectrale et de la corrélation spatiale.

SUMMARY

This paper deals with the joint utilization of spectral and spatial information in performing the segmentation of satellite multispectral images. By taking a Markovian model in a unidimensional simulation, like video scanning, we present some results by which we choose between two types of classification strategies : one using a reduced spatial context and spectral information, performed by feature extraction, another using the entire spectral information without context. The classification is based on the maximum likelihood criterion in all the cases studied. Feature extraction is realized by the Fisher discriminant approach. The simulation results show which strategy to choose in terms of spatial and spectral correlations.

1. Introduction

La segmentation automatique des images joue un rôle important en télédétection par satellite. Cette tâche est souvent réalisée par l'analyse des signatures spectrales des pixels, qui procède par classification multispectrale en utilisant les valeurs numériques des divers canaux d'observation du satellite. En complément, la haute résolution des données permet une analyse spatiale des images.

L'utilisation conjointe de l'information spectrale et de l'information spatiale présente alors l'intérêt d'exploiter pleinement les possibilités du capteur.

Il existe plusieurs approches de l'analyse de l'information spatiale, l'approche texturale a été adoptée avec succès sur des images du satellite SPOT

[1]. L'approche contextuelle, faisant l'objet du présent article, a été abordée par l'étude comparative sur des images multispectrales simulées dans le cas unidimensionnel, type balayage vidéo. Dans un premier temps nous cherchons à quantifier l'apport du contexte, d'une part, et mesurer la dégradation des performances lorsqu'on utilise une stratégie d'extraction de primitives, d'autre part. Dans un deuxième temps, nous cherchons à déterminer des critères permettant de savoir si, dans le cas où le temps de calcul ne permet pas de garder simultanément l'information spectrale complète et le contexte, l'on doit plutôt opérer une extraction de primitives et garder le contexte ou garder l'information spectrale complète en renonçant à l'apport du contexte.



2. Modèle unidimensionnel de l'image

On considère le modèle suivant :

2.1 L'image initiale (La "vérité-terrain" en imagerie satellitaire) qui n'est pas accessible à l'observation, est modélisée par une chaîne de Markov stationnaire du premier ordre, chaque $\xi_n, n=1,2,\dots,N$ étant à valeurs dans l'ensemble de K classes possibles, $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$. La chaîne de Markov est caractérisée par une distribution initiale $\Pi_j = P(\xi_1 = \omega_j)$, $\omega_j \in \Omega$ et une matrice de transition de terme général : $\Pi_{i/j} = P(\xi_n = \omega_i | \xi_{n-1} = \omega_j)$, $\omega_i, \omega_j \in \Omega^2$.

2.2 L'image multispectrale observée est modélisée par une suite de vecteurs aléatoires \mathbf{x}_n , $n=1,2,\dots,N$ où $\mathbf{x}_n = (\mathbf{x}_{n1}, \mathbf{x}_{n2}, \dots, \mathbf{x}_{nm})^T$. Conditionnellement à la chaîne ξ_n ($1 \leq n \leq N$), la suite \mathbf{x}_n ($1 \leq n \leq N$) est un processus gaussien markovien stationnaire, caractérisé par K^2 distributions du couple $(\mathbf{x}_{n-1}, \mathbf{x}_n)$ conditionnelles à (ξ_{n-1}, ξ_n) , soit :

$$p_{ij}(\mathbf{y}_n) = P(\mathbf{y}_n | \xi_{n-1} = \omega_i, \xi_n = \omega_j) \quad i, j = 1, 2, \dots, K$$

avec $\mathbf{y}_n = \begin{pmatrix} \mathbf{x}_{n-1} \\ \mathbf{x}_n \end{pmatrix}$. On notera la moyenne $\mathbf{u}_{ij} = \begin{pmatrix} \mathbf{u}_i \\ \mathbf{u}_j \end{pmatrix}$, et la matrice de covariance, que l'on suppose identique, $\mathbf{c}_{ij} = \mathbf{c}$, $i, j = 1, 2, \dots, K$ avec $\mathbf{c} = \begin{pmatrix} \mathbf{c}_1 & \mathbf{c}_2 \\ \mathbf{c}_2 & \mathbf{c}_1 \end{pmatrix}$ où

$$\begin{aligned} \mathbf{c}_1 &= E[(\mathbf{x}_n - \mathbf{u}_k)(\mathbf{x}_n - \mathbf{u}_k)^T | \xi_n = \omega_k] \\ \mathbf{c}_2 &= E[(\mathbf{x}_{n-1} - \mathbf{u}_i)(\mathbf{x}_n - \mathbf{u}_j)^T | \xi_{n-1} = \omega_i, \xi_n = \omega_j] \end{aligned} \quad i, j, k = 1, 2, \dots, K$$

3. Extraction de primitives

L'augmentation de l'information spatiale, utilisée dans la stratégie contextuelle, implique un accroissement considérable du temps de calcul. L'étape d'extraction de primitives a pour but de réduire la dimension de l'espace des observations. La méthode retenue, basée sur les fonctions discriminantes de Fisher, se formalise de la manière suivante [2] :

Soit $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_D)^T$ le vecteur de primitives originales d'un pixel de l'image. On cherche à trouver des combinaisons linéaires de \mathbf{y} , c'est à dire un matrice \mathbf{a} rectangulaire de dimension $D \times d$ avec

$d < D$, à l'aide d'un critère \mathbf{J} tel que

$$\mathbf{J}(\mathbf{a}^*) = \max_{\mathbf{a}} \mathbf{J}(\mathbf{a})$$

Nous utilisons ici les fonctions discriminantes de Fisher pour $\mathbf{J}(\mathbf{a})$, soit

$$\mathbf{J}(\mathbf{a}) = \frac{|\mathbf{a}^T \mathbf{s}_B \mathbf{a}|}{|\mathbf{a}^T \mathbf{s}_W \mathbf{a}|}$$

où

$$\mathbf{s}_B = \sum_{i=1}^K \Pi_i (\mathbf{u}_i - \mathbf{u})(\mathbf{u}_i - \mathbf{u})^T \quad \text{matrice de covariance interclasse}$$

$$\mathbf{s}_W = \sum_{i=1}^K \Pi_i \mathbf{c}_i \quad \text{matrice de covariance intraclasse}$$

avec

$$\mathbf{u} = \sum_{i=1}^K \Pi_i \mathbf{u}_i$$

\mathbf{u}_i la moyenne de la i ème classe

\mathbf{c}_i la matrice de covariance de la i ème classe

Π_i la probabilité a priori de la i ème classe

Les vecteurs colonne de \mathbf{a}^* sont alors les vecteurs propres de $\mathbf{s}_W^{-1} \mathbf{s}_B$ rangés par ordre décroissant des valeurs propres correspondantes.

Les primitives extraites sont ensuite données

par

$$\mathbf{y}' = \mathbf{a}^{*T} \mathbf{y}$$

avec

$$\mathbf{y}' = (y'_1, y'_2, \dots, y'_d)^T$$

4. Algorithme de segmentation

Rappelons d'abord la méthode d'estimation par maximum de vraisemblance a posteriori [3] :

$$\xi_n = \omega_i \quad \text{si} \quad L_i(\mathbf{v}_n) = \max_j L_j(\mathbf{v}_n) \quad j = 1, 2, \dots, K$$

où $L_j(\mathbf{v}_n) = P(\xi_n = \omega_j, \mathbf{v}_n)$ est la probabilité conjointe de la classe ω_j et du vecteur de primitives \mathbf{v}_n qui représente le pixel n .

Dans le cas classique (non-contextuel), $\mathbf{v}_n = \mathbf{x}_n$, on ne tient compte que de l'information spectrale, donc

$$\begin{aligned} L_j(\mathbf{v}_n) &= P(\xi_n = \omega_j, \mathbf{x}_n) \\ &= P(\xi_n = \omega_j) P(\mathbf{x}_n | \xi_n = \omega_j) \\ &= \Pi_j p_j(\mathbf{x}_n) \end{aligned}$$

Dans le cas contextuel, $\mathbf{v}_n = \mathbf{y}_n = \begin{pmatrix} \mathbf{x}_{n-1} \\ \mathbf{x}_n \end{pmatrix}$, on tient compte, pour classer le pixel n , de l'information spatiale apportée par l'observation \mathbf{x}_{n-1} . La probabilité conjointe de (ξ_n, \mathbf{y}_n) est donnée par :

$$\begin{aligned}
 L_j(\mathbf{y}_n) &= \sum_{i=1}^K P(\xi_{n-1} = \omega_i, \xi_n = \omega_j, \mathbf{y}_n) \\
 &= \sum_{i=1}^K P(\xi_{n-1} = \omega_i, \xi_n = \omega_j) \\
 &\quad \times P(\mathbf{y}_n | \xi_{n-1} = \omega_i, \xi_n = \omega_j) \\
 &= \sum_{i=1}^K \Pi_{ij} p_{ij}(\mathbf{y}_n)
 \end{aligned}$$

5. Simulations

Dans toutes les simulations, on se restreint au cas de trois classes (K=3) avec les probabilités a priori $\Pi_1 = 0,45, \Pi_2 = 0,35,$ et $\Pi_3 = 0,20.$ La loi de (ξ_n) est donc définie dès qu'on se donne la matrice de transition : les simulations portent sur deux cas (indépendant et dépendant) suivants :

a. cas indépendant

$$\Pi_{i/j} = \begin{pmatrix} 0.450 & 0.450 & 0.450 \\ 0.350 & 0.350 & 0.350 \\ 0.200 & 0.200 & 0.200 \end{pmatrix}$$

b. cas dépendant

$$\Pi_{i/j} = \begin{pmatrix} 0.350 & 0.579 & 0.450 \\ 0.450 & 0.121 & 0.525 \\ 0.200 & 0.300 & 0.025 \end{pmatrix}$$

A l'image initiale, préalablement convertie en image multispectrale de deux canaux aux valeurs $\mathbf{u}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ est ajouté le bruit gaussien (\mathbf{z}_n) de moyenne nulle. La corrélation entre \mathbf{z}_{n-1} et \mathbf{z}_n est définie par $\mathbf{c} = \begin{pmatrix} \mathbf{c}_1 & \mathbf{c}_2 \\ \mathbf{c}_2 & \mathbf{c}_1 \end{pmatrix}$ avec

$$\mathbf{c}_1 = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{c}_2 = \begin{pmatrix} \delta & 0 \\ 0 & \delta \end{pmatrix}$$

où ρ (la corrélation spectrale entre les deux canaux) et

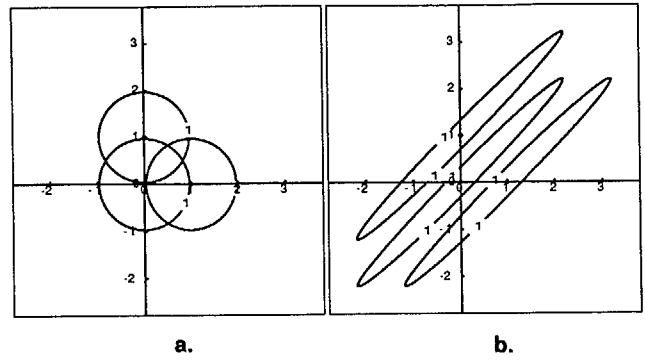


Fig. 1 Contour du niveau 0,1 des répartition gaussiennes (a) $\rho = 0,00$ (b) $\rho = 0,99$

δ (la corrélation spatiale entre les observations enregistrées par un même canal) sont les paramètres à faire varier ($0 \leq \rho, \delta < 1$ et $\rho + \delta < 1$). La figure 1 montre le contour du niveau 0,1 des répartition gaussiennes.

Pour le cas contextuel, la phase d'extraction de primitives réduit la dimension de l'espace de classification de quatre (dimension du vecteur \mathbf{y}) à deux (dimension du vecteur de primitives dans le cas non-contextuel).

6. Interprétation des résultats

6.1 Notations

- S_1 : Stratégie contextuelle utilisant les primitives originales
- S_2 : Stratégie contextuelle utilisant les primitives extraites
- S_3 : Stratégie non-contextuelle utilisant les primitives originales
- C_1 : Cas indépendant
- C_2 : Cas dépendant

6.2 Remarques générales

- La stratégie S_1 est, dans tous les cas, au moins aussi efficace que S_2 et S_3 .

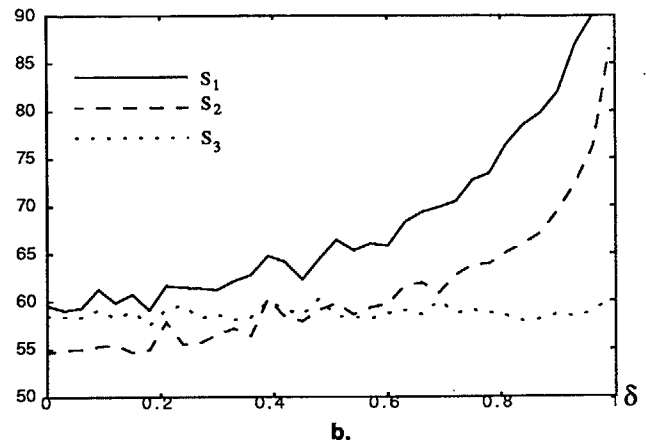
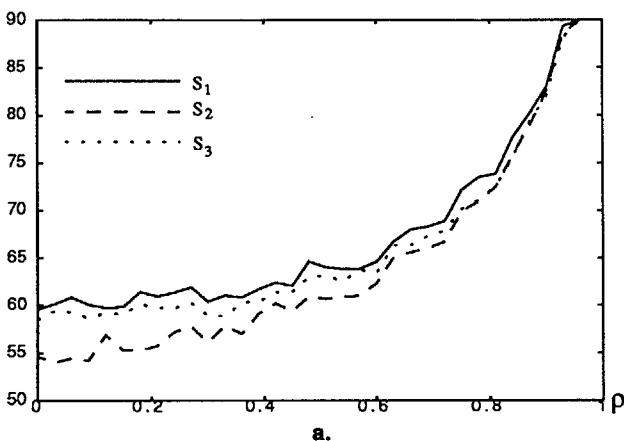


Fig. 2 Les courbes des figures a. et b. visualisent le comportement des diverses stratégies, en termes du taux de bon classement, dans le cas C2, en fonction de ρ (pour $\delta = 0$) et en fonction de δ (pour $\rho = 0$) respectivement.



- A ρ constant la croissance de δ augmente l'efficacité des stratégies contextuelles, même dans le cas C_1 (Fig.3 a, b, d, e), par contre elle n'influe pas sur l'efficacité des stratégies non-contextuelles (Fig.3 c, f).
- A δ constant l'efficacité de toutes les stratégies croît avec ρ . Ce fait s'explique par la disposition spatiale des moyennes (fig. 1). Une autre répartition des moyennes pourrait donner des résultats très différents.

6.3 Choix entre S_2 et S_3

- Le choix de S_3 paraît plus judicieux pour $0 \leq \rho + \delta \leq 0,67$ dans le cas C_1 et pour $0 \leq \rho + \delta \leq 0,5$ dans le cas C_2 .
- Dans le cas C_1 et pour $0,67 \leq \rho + \delta \leq 1,0$ les efficacités de S_2 et S_3 sont comparables sur l'ensemble $0,5 \leq \rho < 1,0$: on peut, dans ce cas, opter pour S_3 dont la mise en oeuvre est plus simple.
- Comme ci-dessus, on pourra opter pour S_3 sur l'ensemble $\{\rho + \delta \geq 0,5\} \cap \{\rho \geq 0,5\}$ dans le cas C_2 .

7. Conclusion

Cette comparaison de stratégies de classification a été conduite sur des images de simulation. Elle donne des indications claires relativement au choix possible entre stratégie avec information contextuelle et spectrale extraite et stratégie non-contextuelle avec information spectrale complète, en fonction des corrélations spatiale et spectrale.

Cependant, avant de généraliser la portée de ces résultats, il conviendra de compléter cette étude. Une application de ces méthodes à des images du satellite SPOT est en cours.

Références

- [1] Lejeune J.C., Bonn F., Roux C., and Odoux B., "Apport de l'information de texture à la segmentation des images du satellite SPOT en milieu urbain", in *Montech'87 (Montreal)*, 10-12 Nov. 1987.
- [2] Duda R.O., and Hart P.E., *Pattern Classification and Scene Analysis*, New York : Wiley, 1973.
- [3] Devijver P.A., and Kittler J., *Pattern Recognition : A Statistical Approach*, London : Prentice-Hall, 1982.

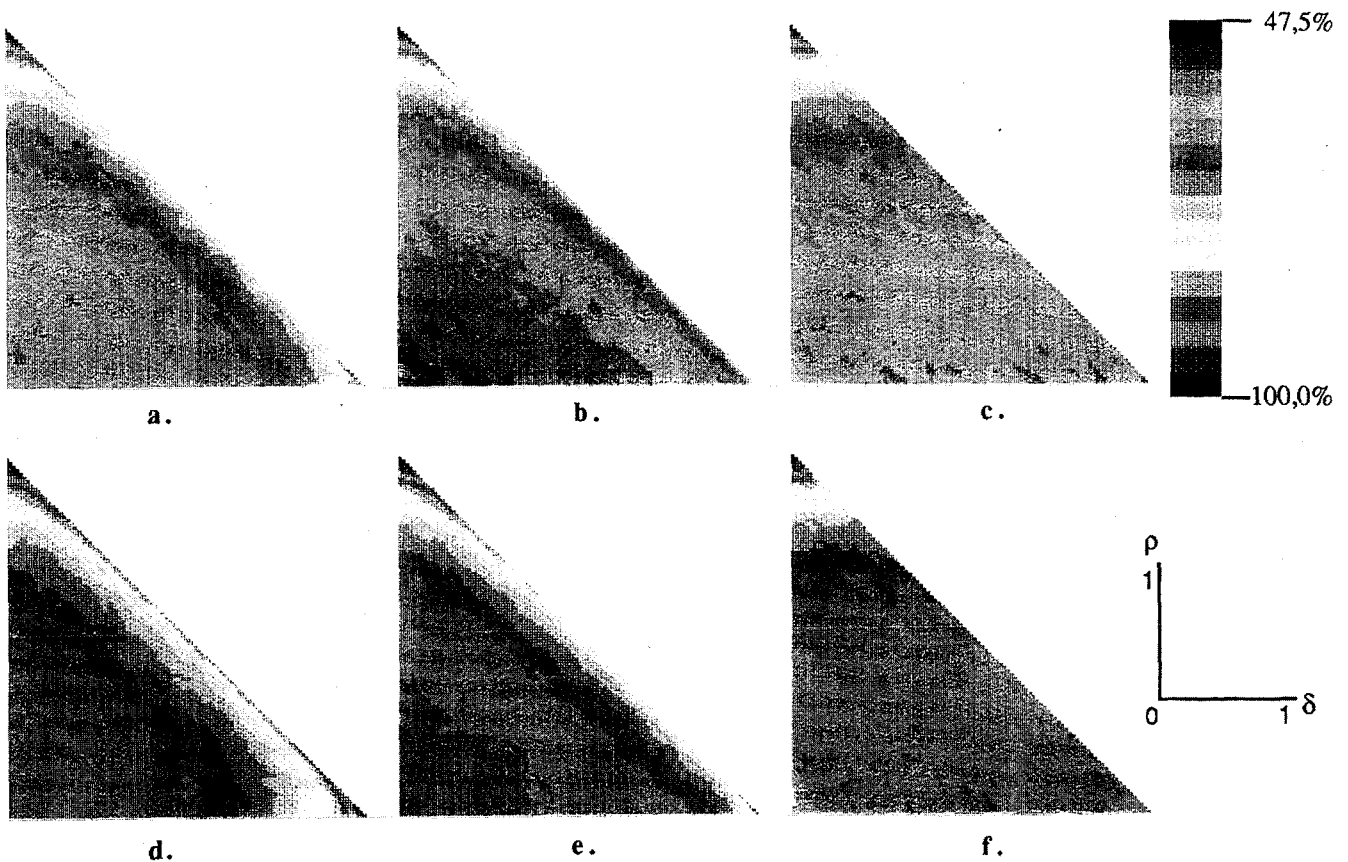


Fig. 3 Les taux de bon classement, dans le cas indépendant (a, b, c) et dans le cas dépendant (d, e, f), obtenus par la stratégie contextuelle utilisant les primitives originales (a, d), utilisant les primitives extraites (b, e), et par la stratégie non-contextuelle (c, f).