

Essai d'estimation de cultures annuelles à partir d'une nomenclature établie sur une image satellitaire.

N'KANZA S.

LATES-C.A.M.S. - E.H.E.S.S. 54, Bd Raspail 75006 PARIS. (FRANCE)

Tél.: 49-54-21-69/49-54-20-36

Télécopieur: 45-44-93-11. Téléex :MSH 203104 F

Résumé.

Nous disposons d'une image SPOT classée, obtenue par combinaison d'une régression "par boules", d'un seuillage, d'un écart-type local, de l'analyse discriminante et des opérateurs de texture de morphologie mathématique.

Sur cette image, nous tirons un échantillon constitué des segments carrés de 720mx720m, sur lesquels nous définissons des variables cultures (blé par exemple) et en estimons la superficie par des méthodes de statistique spatiale.

1°) Introduction.

Nous cherchons à évaluer la superficie de certaines catégories de cultures et d'en faire un inventaire à partir des segments géométriques tirés sur des données SPOT.

Mais avant d'obtenir cette classification, nous avons une chaîne de plusieurs traitements.

Dans cet article, nous exposons seulement les grandes lignes méthodologiques d'un des traitements qui est la régression "par boules" [5] et relatons brièvement les autres méthodes qui nous ont permis d'aboutir à la nomenclature établie sur les images. Les estimations par des méthodes de statistique spatiale sont l'objet d'un dernier volet d'une thèse de doctorat en voie d'achèvement [6]. Leurs applications seront exposées dans une publication ultérieure.

Nous disposons au départ de trois canaux XS1, XS2 et XS3 provenant de deux scènes (39-252 HRV1 et 40-252 HRV2) du 1^{er} Mai 1986. La zone de travail porte uniquement sur la deuxième scène et concerne plus de la moitié du département de l'Essonne situé au sud de Paris.

2°) Régression "par boules".

La corrélation entre XS1 et XS2 étant très élevée, nous avons créé par la méthode de régression "par boules", un néo-canal issu des deux premiers. En effet, dans le traitement d'images satellitaires, il est fréquent que certains canaux (par exemple XS1 et XS2 de SPOT ou MSS4 et MSS5) soient fortement corrélés.

Et, il est d'usage courant, au vu de leur histogramme, donc de la dynamique de leurs valeurs radiométriques, de privilégier un canal au détriment d'un autre. Les raisons souvent évoquées sont entre autres, la faible dynamique d'un canal qui rend son exploitation sans grand intérêt, la "catégorie d'occupation du sol" étudiée est plus prépondérante dans un canal que dans un autre, etc... Toutes ces raisons peuvent avoir leur justification thématique. Mais, il n'en demeure pas moins que le rejet, souvent hâtif ou abusif, d'un des deux canaux est courant dans les applications. C'est pour cette raison que nous avons proposé l'estimateur de régression "par boules" qui permet de créer à partir de deux canaux fortement corrélés et éventuellement bruités, un néo-canal filtré du bruit.

2-a) Le modèle mathématique.

Le modèle mathématique dont nous nous sommes inspirés pour proposer l'estimateur de régression "par boules" est le suivant:

$$Z(x) = Y(x) + \varepsilon(x) \quad (1)$$

$Z(x)$: variable observée au pixel x .

$Y(x)$: le signal au pixel x .

$\varepsilon(x)$: le bruit au pixel x .

Ce modèle est monovarié [4] et utilise comme estimateur:

$$\hat{Y}(x) = E[Y(x)] + C_{YZ}^{-1} (Z(x) - E[Z(x)]) \quad (2)$$

Dans notre cas, nous adoptons le modèle bivariable suivant:

$$Z_1(x) = Y(x) + \varepsilon_1(x), \quad i = 1, 2 \quad (3)$$

qui assigne la même valeur à la variable sous jacente Y. Et nous faisons les hypothèses suivantes:

a) $E[\varepsilon_1(x)] = 0, \forall i, \forall x.$

Le bruit est de moyenne nulle dans un voisinage donné du pixel x.

b) $Cov[\varepsilon_1(x), \varepsilon_j(x)] = 0, \forall i \neq j, \forall x \neq y.$

Le bruit observé sur un pixel x est indépendant de celui d'un autre pixel y.

De a) il vient, pour tout voisinage d'un pixel x

c) $E[Z_1(x)] = E[Z_2(x)] = E[Y(x)] = m(x)$

Cette conséquence exprime le fait que sur des petits voisinages, la moyenne locale est supposée la même pour les deux canaux. l'estimateur de régression "par boules" est un filtre local linéaire, calculé dans un voisinage donné de chaque pixel x. Il s'écrit:

$$\hat{Y}(x) = E[Y(x)] + 1/2\lambda_x [(Z_1 + Z_2)(x) - (m_1 + m_2)(x)]$$

$\forall x.$ (4)

avec $m_i = E[Z_i]$, $1 \geq \lambda_x \geq \frac{|Cov[Z_1(x), Z_2(x)]|}{\sqrt{Var[Z_1(x)] Var[Z_2(x)]}} \geq 0$

Le pondérateur λ_x est la valeur absolue du coefficient de corrélation local.

Si n est le nombre total de pixels sur chaque image, nous avons

$$\varepsilon(x_i) = Z(x_i) - \hat{Y}(x_i), \quad \forall i = 1, \dots, n$$

qui est l'erreur d'estimation au pixel x_i . Et nous appellons le rapport du signal sur le bruit au pixel la valeur

$$R_{x_i} = \frac{Var[\hat{Y}(x_i)]}{Var[\varepsilon(x_i)]}$$

Le rapport moyen du signal sur le bruit est exprimé par

$$R = \frac{1}{n} \sum_{i=1}^n R_{x_i} \quad (5)$$

3°)

A l'issue de la régression "par boules", nous disposons de deux canaux (le néo-canal et le troisième canal) sur lesquels nous avons appliqué un seuillage, un écart-type local dans un voisinage glissant de taille impaire [3], une analyse discriminante et des filtres de morphologie mathématique qui sont l'érosion et la dilatation avec un élément structurant de connexité 4 [1] pour coder les pixels qui appartiennent à un intervalle donné, réhausser les contours et réaliser une partition plus fine de l'espace agricole pour obtenir notre nomenclature.

4°) Applications.

-Corrélation(XS1, XS2) = 0.9592

-Rapport moyen du signal sur le bruit en fonction du voisinage choisi

	3x3	5x5	7x7
R	8.588	7.146	5.370

-Quelques seuils utilisés

néo-canal	53 ≤ x ≤ 65	} espace urbain
canal 3	40 ≤ x ≤ 65	
néo-canal	39 ≤ x ≤ 53	} forêt
canal 3	45 ≤ x ≤ 68	

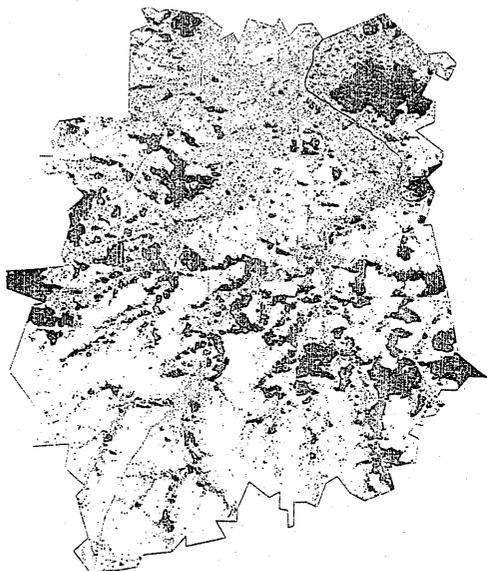
-Voisinage utilisé pour l'écart-type local = 3.

- Eléments structurants de connexité 4

érosion			dilatation		
2	0	2	2	1	2
0	1	0	1	0	1
2	0	2	2	1	2

-Image classée

DEPARTEMENT ESSONNE
Masque Forêt (noir), Urbain (gris clair),
Eau (gris foncé), Esp. Agricole (blanc)



- Nomenclature obtenue

Code	Désignation
1	Urbain
2	Forêt
3	Eau
105	Céréales de printemps
106	Céréales d'hiver
107	Autres cultures de printemps
108	Prairies
109	Colza

5°) Conclusion.

La corrélation entre les deux premiers canaux est très élevée. A partir du rapport moyen du signal sur le bruit, nous avons retenu une fenêtre de taille 3 pour la régression par "boules" et pour l'écart-type local. C'est avec cette taille que nous perdons le moins de structures sur les images. La nomenclature obtenue a été jugée bonne, après confrontation avec les enquêtes terrain. Les estimations par des méthodes de statistique spatiale seront faites sur les 504 segments tirés à partir de l'image classée.

6°) Bibliographie.

- [1] BONAN, N. " Manuel des utilisateurs du Laboratoire de Télédétection" LATES - C.A.M.S. E.H.E.S.S. 54 ,Bd Raspail 75006, Paris.
- [2] FOURNIER, Ph. ; GILG, J.P.; GEROYANNIS, H. "Evaluation des surfaces des cultures annuelles et inventaire du territoire dans le bassin parisien par combinaison des enquêtes statistiques classiques et des données SPOT (PEPS 127). SPOT1. Utilisation des images Bilan. Résultats. Paris, Novembre 1987.
- [3] GEROYANNIS, H. " Manuel des utilisateurs du Laboratoire de Télédétection" LATES - C.A.M.S. E.H.E.S.S. 54 ,Bd Raspail 75006, Paris.
- [4] KUAN, D.T. et al "Adaptive noise smoothing filter for images with signal-dependent noise" IEEE transactions on pattern analysis and machine intelligence. Vol PAMI-7, n° 2, March 1985.
- [5] N'KANZA, S. "Comparaison de trois méthodes statistiques appliquées à des images de microsonde". Mastère Spécialisé de Géostatistique. Centre de Géostatistique E.N.S.M.P. . Fontainebleau, Septembre 1988.
- [6] N'KANZA, S. " Géostatistique. Etude de la structure d'une image. Optimisation d'un plan de sondage". Thèse en fin de rédaction. L.A.T.E.S. C.A.M.S. - E.H.E.S.S. Paris