

CLASSIFICATION PIXELS ET CLASSIFICATION PAR ZONES : TESTS SUR IMAGE SPOT EN MILIEU URBAIN

WEBER C.¹, HIRSCH J.¹, SERRADJ A.²

RESUME

L'analyse du milieu urbain à des fins de modélisation, se caractérise en traitement d'image par la nécessité de maîtriser l'hétérogénéité du milieu. Cette maîtrise passe par une certaine généralisation de l'information satellitaire obtenue. Dans cette optique, on est en droit de se demander quelle approche de l'analyse pixel ou de l'analyse zonale est la plus intéressante. Cette présentation fait état de tests sur l'analyse zonale, et des conclusions qui peuvent être tirées en comparaison avec l'approche plus usuelle par traitement pixel.

Le but de la démarche est ici, d'identifier un niveau spatial, qualifié de "mésos-niveau", qui permette d'améliorer les analyses de modélisation de la structure du bâti et éventuellement de la population en fonction de l'information satellitaire. Les analyses menées précédemment ont permis de soulever différents problèmes inhérents à l'hétérogénéité du milieu urbain, et dont les conséquences en termes de modélisation peuvent être préjudiciables. Le traitement par pixels conduit en effet à des erreurs de classification pour certains "objets" urbains. La discrimination d'objets tels des berges bétonnées ou empierrées, ou encore d'infrastructures ne peut se faire facilement même en utilisant une classification supervisée du type analyse discriminante (BMDP). La variabilité des espaces intra-urbains présentant des sols cultivés, ou nus, des cimetières, ou encore des zones jouxtant un habitat caractérisé par des collectifs (parkings, patio...) biaise les résultats d'analyses spatiales ultérieures. L'identification d'un "mésos-niveau" bien qu'engendrant une généralisation de l'information satellitaire semble être une approche intéressante dans ce cas. Il nous faut donc vérifier si la généralisation de l'information introduite, permet de préserver une certaine spécificité des réponses spectrales au niveau des zones considérées. C'est pourquoi il est nécessaire de définir des entités aréales à une échelle suffisamment fine, pour que l'agrégation de l'information de départ ne subisse pas trop d'altération. L'introduction d'une composante spatiale par l'intermédiaire d'un fond numérisé peut permettre de localiser et de caractériser des arrangements spatiaux qui peuvent par la suite être utilisés à des fins de croisement de données multisources.

Les données utilisées sont d'une part une Image SPOT XS de juin 1986, une Image panchromatique de Septembre 1986, et d'autre part un fond numérisé des îlots

¹URA 902 CNRS / GSTS - Strasbourg.

²Equipe COSTEL - Rennes.

INSEE de Strasbourg, référence 1982 selon une projection locale (Sausheim), et les données de recensement correspondantes.

Plusieurs types de traitements ont été appliqués sur deux groupes de données:

- d'une part les données radiométriques brutes.
- d'autre part les informations sur l'occupation du sol à partir d'une classification supervisée (en pourcentages par entités aréales)

Données radiométriques

Analyse factorielle zonale

Analyse de classification
non supervisée

Analyse de classification
supervisée

Données occupation du sol

Analyse factorielle zonale

Analyse de classification
non supervisée

Analyse de classification
supervisée

Analyse à partir du maximum

Si le deuxième groupe de variables ne pose pas de problème, puisqu'il s'agit des données traitées, le premier nous a amené à nous pencher sur la justification des variables radiométriques à retenir afin de caractériser au mieux les zones retenues.

I - ANALYSES DES VARIABLES RADIOMETRIQUES

Différents tests ont été faits d'une part pour évaluer la redondance d'information contenue dans certaines mesures statistiques usuelles (moyenne, écart-type, coefficients de variation, amplitude), et d'autre part la stabilité des résultats en fonction des choix successifs de variables introduites dans les analyses. Bien entendu certains résultats ne sont pas étonnant compte tenu de la corrélation existant entre les canaux XS1 et XS2.

Les corrélations obtenues semblent confirmer cette analyse, permettant ainsi de déterminer au mieux les variables nécessaires aux analyses zonales.

	A1	M1	Ec1	A2	M2	Ec2	A3
A 1							
M 1							
Ec 1	.871						
A 2	.961		.834				
M 2		.939					
Ec 2	.836		.939	.877			
A 3	.687		.544	.745		.636	
M 3							
Ec 3	.556		.505	.616		.601	.914
A 4	.812		.717	.787		.692	.577
M 4		.885			.877		
Ec 4	.698		.807	.662		.745	

Matrice des corrélations les plus élevées.

A 1, A 2, A 3 correspondent aux amplitudes des données sur les différents canaux XS. M1, M2, M3, sont les valeurs des moyennes sur XS 1, XS2, XS3. Ec 1, Ec 2, Ec 3 en sont les écarts-types. De même pour le panchromatique noté A4, M4, Ec 4.

Ces résultats démontrent bien les fortes relations existant entre les différentes variables. Dans un tel cas de figure il semble acceptable de se contenter d'un nombre plus faible de variables. Il est intéressant de noter la stabilité des facteurs obtenus dans les analyses à 12, 6 et 4 variables, avec ou sans rotation des axes, où seules les variables les plus discriminantes sont conservées. En l'occurrence dans la dernière analyse, seules les moyennes des XS2 et XS3 et leurs écart-types sont conservés (variance expliquée: 90%).

1° facteur: variabilité du signal sur les canaux XS1, XS2 et P.

2° facteur: moyenne du signal sur les mêmes canaux.

3° facteur: moyenne du signal sur XS3.

4° facteur: variabilité du signal sur XS3.

Il faut noter que la composition colorée des scores factoriels zonaux donne des résultats interprétatifs tout à fait intéressants, en termes d'analyse urbaine; des zones de différenciation délicate comme certaines zones pavillonnaires ou portuaires apparaissent à ce stade plus judicieusement caractérisées.

Une classification non supervisée (K-Means BMDP) a été tentée, dix classes ont été choisies correspondant au nombre de catégories d'occupation du sol identifiables à Strasbourg dans le tissu urbain en termes d'occupation du sol.

	M 2	Ec2	M 3	Ec3
Between	8631.38	961.96	48093.672	2718.75
Within	12.67	6.30	13.573	9.67
F-ratio	680.87	152.58	3543.452	281.07

Analyse de variance: (D.L. 9, 2806).

Au vue de ces résultats il semblerait que la variable la plus discriminante soit la moyenne sur XS3, les groupes ainsi différenciés semblent statistiquement satisfaisants; cependant les résultats cartographiques de cette classification sont décevants en termes d'interprétation spatiale. Si quelques arrangements spatiaux se dégagent comme le centre-ville par exemple ou les emprises industrielles, beaucoup de mauvaises affectations peuvent être notées pour certaines entités aréales corroborant les difficultés de traitements de l'analyse "pixel".

Une classification supervisée a donc été tentée en tenant compte de zones d'entraînement "étalon" choisies pour leurs caractéristiques radiométriques, et notre connaissance du terrain. Quatorze groupes d'échantillons test ont été définis:

- Trois groupes caractérisant le bâti dense continu, regroupant des structures bâties de différents types et âges.
- Un groupe qualifiant les vieux noyaux suburbains des quartiers ou des communes environnantes.
- Trois groupes caractérisant les ensembles collectifs, tenant compte de leur structure spatiale et de leurs caractéristiques urbanistiques.
- Deux groupes de bâti discontinu, pavillonnaire.

- Deux groupes de grandes emprises industrielles ou ferroviaires.
- Deux groupes caractérisant la végétation.
- Un groupe de surfaces en eau.

Les pourcentages de classification correcte varient entre 50 et 82 % pour le bâti à usage résidentiel dominant (sauf pour un groupe spécifique: Ville Allemande), entre 60 et 80% pour le bâti industriel et 100% pour les aires végétales et d'eau. Une bonne discrétisation s'établit au sein du schéma des variables canoniques entre ces quatre types d'occupation du sol.

Le groupe "Ville Allemande" a été testé pour la spécificité de la structure spatiale de cette zone urbaine à Strasbourg, en effet ce quartier se différencie très bien dans le tissu morphologique de la ville. Il semble cependant qu'il n'y ait pas spectralement de différenciation suffisamment explicite, pour vouloir en faire un groupe particulier.

Nous avons mené la même analyse en n'utilisant que les moyennes comme variables, pour tester la part de la variabilité dans un tel contexte. Il semble évident que cette notion de variabilité soit nécessaire dans l'analyse, certains sous-espaces urbains ne se définissant que par cette dimension. C'est le cas par exemple des zones portuaires à l'est de la ville où la présence de vieux bâtiments industriels, d'eau, de quais et de dépôts de tous genre entraîne une hétérogénéité non contournable.

Types	Groupes	Classes	M2, Ec2, M3, Ec3	M2, M3
Bâti Dense	Dense 1	1	50.	62.5
	Dense 2	2	66.7	66.7
	Ville Allemande	3	11.1*	22.2**
	Noyaux	4	82.4	70.6
Eau	Eau	5	100.	50.
Bâti Collectif	Haute pierre	6	75.	75.
	Canardière	7	66.7	11.1
	Campus	8	60.	60.
Bâti Pavillonnaire	Meinau	9	75.	58.3
	Neudorf	10	54.5	36.4
Industriel	Plaine des Bouchers	11	80	80
	Vieil Industriel	12	60	0
Végétation	Végétation mixte	13	100	100
	Végétation	14	100	75

* Ce groupe se scinde surtout sur différents types de bâtis.

** La répartition est ici plus confuse, des affectations se faisant aussi bien dans le bâti dense que dans le vieil industriel ou la végétation mixte.

Ce tableau croisé des pourcentages de classification, présente des résultats spectaculaires en ce qui concerne en particulier les classes de vieil industriel et d'un type de bâti collectif. Il semble donc tout à fait nécessaire de conserver la variabilité dans une approche zonale, l'utilisation des seules moyennes de canaux XS ne définissant de façon satisfaisante que des zones plutôt homogènes. La cartographie des affectations permet d'individualiser des arrangements spatiaux identifiables et utilisables en analyse spatiale. Cependant des aberrations sont décelables dues à la variabilité du signal dans certaines zones comme les canaux ou les cimetières, ceci est accentué en fonction de la taille des zones numérisées.

II ANALYSE SUR LES POURCENTAGES D'OCCUPATION DU SOL

Cette approche utilise les pourcentages d'occupation du sol obtenus après une analyse de classification supervisée (Analyse discriminante) dont les résultats ont été jugés satisfaisants aux vues des pourcentages de pixels bien classés. Six types d'occupation du sol ont été retenus

- Bâti dense continu, bâti caractérisant les grands ensembles collectifs, bâti pavillonnaire, grandes emprises industrielles, surfaces végétales et surfaces en eau.

Les corrélations obtenues entre ces variables sont faibles et démontrent bien les oppositions entre les différentes structures spatiales du tissu urbain.

L'analyse factorielle en composante principale à quatre facteurs (80% de la variance) caractérise plusieurs types de structures spatiales.

Facteur 1: structures bâties opposant le bâti dense au bâti pavillonnaire discontinu.

Facteur 2: structures caractérisant les ensembles collectifs et les emprises industrielles.

Facteur 3: surfaces végétales.

Facteur 4: surfaces en eau.

La composition colorée des scores factoriels met en évidence, elle aussi, les arrangements spatiaux au sein du tissu urbain. Les différentes structures bâties sont bien différenciées spatialement et typologiquement. Une classification des scores factoriels devrait pouvoir renforcer les résultats obtenus.

Une classification non supervisée, sur dix classes a été faite de la même manière que pour les données radiométriques.

Analyse de variance: (D.L. 9, 2250)

	Dense	Ensembles	Pavillonnaire	Industriel	Végétation	Eau
Between	217502.62	34378.75	151118.87	34652.12	148006.62	10421.31
Within	87.90	81.43	96.50	64.41	83.85	7.00
F-ratio	2474.22	422.18	1565.85	537.92	1765.12	1487.33

Les groupes sont, comme précédemment, statistiquement satisfaisants; cette discrimination entraîne des résultats cartographiques aisés à interpréter, où des arrangements spatiaux s'individualisent en accord avec les structures spatiales reconnues sur le terrain.

De la même façon que pour les analyses sur données radiométriques, une analyse discriminante a été menée sur les pourcentages en tenant compte des "étalons" prédéterminés. Si les pourcentages de classification sont en général plutôt élevés, les deux types identifiés auparavant pour leur hétérogénéité présentent, dans ce cas aussi, des résultats faibles.

Types	Groupes	Classes	M2, Ec2, M3, Ec3	%Occupation du sol
Bâti Dense	Dense 1	1	50.	75.
	Dense 2	2	66.7	66.7
	Ville Allemande	3	11.1	22.2
	Noyaux	4	82.4	35.3
Eau	Eau	5	100.	100.
Bâti Collectif	HautePierre	6	75.	87.5
	Canardière	7	66.7	44.4
	Campus	8	60.	80.
Bâti Pavillonnaire	Meinau	9	75.	83.3
	Neudorf	10	54.5	72.7
Industriel	Plaine des Bouchers	11	80	100
	Vieil Industriel	12	60	30
Végétation	Végétation mixte	13	100	100
	Végétation	14	100	100

De plus la représentation des affectations met en évidence quelques aberrations notamment en termes de couverture végétale ou des zones périphériques qui correspondent à des zones d'habitat épars et qui dans ce cas de figure sont affectées au groupe "habitat collectif".

CONCLUSION

Dans cette expérimentation nous avons tenté de déterminer si un "meso-niveau" pouvait être défini en utilisant une approche zonale pour les traitements d'image. Selon les différents résultats obtenus nous pouvons penser qu'une telle approche est envisageable mais sous certaines réserves.

En effet nous avons utilisé comme définition de zonage, le fond numérisé à l'îlot dont nous disposons, or à ce stade il est pertinent de se demander si ce choix est judicieux et s'il ne faut pas pousser les recherches vers les possibilités de définition de zonage uniquement en fonction de l'image. Les contraintes identifiées reposent en particulier sur les effets de taille introduits par des entités spatiales disparates notamment entre le centre et la périphérie. Ces disparités entraînent un accroissement de l'hétérogénéité qui perturbe les résultats d'analyse. Un des axes de recherche ultérieurs serait de s'affranchir d'un tel découpage, par une approche de filtrage zonal défini en termes de texture par exemple ou l'introduction d'un carroyage.

Pour pallier les problèmes d'hétérogénéité par zone, une seconde approche pourrait être envisagée, celle utilisant les maxima seuillés des pourcentages d'occupation du sol partant du principe d'acceptation de la classification par pixels. Des types de signatures spatiales ou arrangements spatiaux, pourraient être définies mettant en place la possibilité de règles d'inférence pour l'affectation des entités

BIBLIOGRAPHIE

- J.P. DUCHEMIN & B. LORTIC 1990 - La transformation rapide des espaces urbains: identification, et suivis cartographiques sur images SPOT, ~UITO et Marne-la-Vallée. Villes et Citadins du Tiers Monde, Cahier n°4, ORSTOM.
- A. LEFEUVRE, J. HIRSCH, C. WEBER & P. de FRAIPONT 1990 - Télédétection et SIG au service de l'Environnement: Etude sur le Ried Centre Alsace. Colloque International de la Société Française de Photogrammétrie et de Télédétection et CNES. Strasbourg. 6 au 9 Novembre.
- MICHEL A. 1988 - Stratification de l'Espace Urbain à partir d'images satellitaires pour réaliser un sondage à objectif démographique. Mise au point et évaluation des méthodes d'analyses des images SPOT et LANDSAT-TM en milieu urbain. Thèse de Doctorat. Ecole des Hautes Etudes en Sciences Sociales. Paris. 2 volumes.