

FUSION D'IMAGES RADAR ET OPTRONIQUE EN TELEDETECTION

A. MICHEL, P. CORNIC

Office National d'Etudes et de Recherches Aérospatiales
29, avenue de la Division Leclerc 92320 Châtillon - France

RÉSUMÉ

Cet article se propose de présenter plusieurs méthodes permettant de tirer profit de la complémentarité d'images radar et visible. Quelques techniques de classification seront confrontées par le biais d'un exemple portant sur un couple d'images SAR RAMSES et SPOT panchromatique.

I- Introduction

Cet article présente et compare différentes techniques de fusion d'images optronique et radar préalablement recalées.

Peu d'études portent sur la fusion d'images optronique et radar. L'intégration de données optiques avec des données radar selon la transformée IHS permet d'améliorer la visibilité d'objets cartographiques [WEL88] et de souligner les traits structuraux [YES91]. Une approche intéressante, intégrant des heuristiques pour la détection de ponts et d'agglomérations, à partir d'un couple SIR/SPOT est proposé dans [HOU91].

La suppression de zones occultées (ombres ou nuages) illustrera la complémentarité des capteurs. Un axe de recherche intéressant et peu exploité sera ensuite dégagé : la fusion d'images au niveau pixel par transformée pyramidale.

Nous présenterons enfin une méthode de classification décentralisée (décision prise au niveau de chaque capteur avant la décision finale) et centralisée (une décision unique est prise à partir de l'ensemble des informations disponibles). Les résultats des différentes méthodes seront exposés et commentés, en mettant l'accent sur les points forts et les lacunes de chacune en fonction du type d'occupation du sol.

II- Prétraitements

Caractéristiques des capteurs.

capteurs	RAMSES	SPOT Pa
	aéroporté	satellite
bande	X	0,51 - 0,73
	4 looks	
résolution	5m	10 m
dépression	33°	inconnue
dynamique	16 bits	8 bits

Les images RAMSES ont été acquises par l'ONERA lors d'une campagne aéroportée.

Les prétraitements suivants ont été effectués sur les images optronique et radar :

- recalage (résolution 5 m) sur un référentiel géographique (DFAD) par rotation translation, l'influence du relief sur la déformation des images étant négligeable dans cette région.

- étalement de la dynamique (égalisation de l'histogramme cumulé à 1%).

- segmentation en régions homogènes (méthode ONERA par croissance de régions utilisant un seuil variable).

ABSTRACT

This paper focuses on methods taking advantages of radar and optical sensor complementarity. A few classification techniques are evaluated with SAR RAMSES and panchromatic SPOT images.

- échantillonnage systématique (taux constant : 1/16) de carrés élémentaires (16 x 16 pixels) afin de constituer une base d'apprentissage commune aux classifications.
- calcul de caractéristiques signal et texture (paramètres d'Harralick) pour chaque région.

III Suppression des ombres

La suppression des ombres et surtout des nuages en imagerie satellitaire est généralement résolue par la prise en compte d'une image d'une autre date. Le problème est différent ici, puisqu'il s'agit d'une part de synthétiser une texture et d'autre part d'en trouver automatiquement le modèle à l'aide d'une image radar.

La présence gênante d'ombres tant sur l'image SPOT (Photo 1) que sur l'image radar (Photo 2) de la centrale de Dampierre, nous a conduit à élaborer une méthode supervisée de suppression de ces zones. L'hypothèse principale sous-jacente à cette méthode suppose que deux régions de textures semblables sur l'image SPOT présentent une texture proche dans l'image radar (et réciproquement). Ce qui revient à supposer que le mode de caractérisation des textures ainsi que leurs méthodes de comparaison sont discriminants pour les différents types d'occupation du sol dans les deux modes. La validité de cette hypothèse a été vérifiée pratiquement pour l'optronique sur différentes images SPOT multispectrales.

III.1 Description de l'algorithme

Actuellement, la suppression des ombres est supervisée, l'opérateur devant désigner à l'aide du curseur la région à supprimer. Le processus se déroule ensuite automatiquement : recherche de l'ensemble des régions "radar" interceptant (par superposition) la région "SPOT" à supprimer, puis séquentiellement pour toutes ces intersections :

- extraction de la texture (matrice de cooccurrences de niveaux de gris)
- recherche la région la plus ressemblante (distance entre matrices de cooccurrences),
- détermination de la région "SPOT" interceptante (par superposition),
- synthèse d'une texture par tirage aléatoire sur la matrice de cooccurrences cumulée de cette région,
- application cette texture sur la partie concernée de la région ombrée

III.2 Commentaires.

En appliquant le principe précédemment décrit, nous avons supprimé sur les images panchromatique et radar les zones d'ombre causées par les cheminées de refroidissement (Photo 3). La recherche des régions textuellement proches a du être



limitée aux régions connexes de celle à supprimer. Sans cette contrainte, des régions semblables relativement éloignées pouvaient être sélectionnées. Un changement d'occupation du sol dû au décalage temporel d'enregistrement des images pouvait alors entraîner la synthèse d'une texture impropre. Sans être totalement indiscernables, ces régions synthétisées permettent une classification correcte et présentent une apparence satisfaisante dans le cas de la création de néo-canaux (composantes principales).

IV- Fusion pixel par transformée pyramidale

Le principe de la transformée de HAAR, s'apparente à celui plus récent de transformée en ondelettes, la différence résidant dans le type de fonction utilisée pour décomposer le signal.

IV.1 Décomposition.

La transformée bidimensionnelle de HAAR décompose itérativement le signal en utilisant les 4 masques suivants :

+1	+1	+1	+1	+1	-1	+1	-1
+1	+1	-1	-1	+1	-1	-1	+1

Les résultat sont disposés comme suit :

α	γ	α	γ	γ
β	ϵ	β	ϵ	ϵ

La décomposition et la recomposition s'effectuent à partir des formules suivantes :

Pour (i paire et $i \in [1, nblig/2]$) et (j paire et $\in [1, nbcoll/2]$)

$$\alpha : H(i,j) = I(i,j) + I(i,j+1) + I(i+1,j) + I(i+1,j+1)$$

$$\beta : H(i+nblig/2,j) = I(i,j) + I(i,j+1) - I(i+1,j) - I(i+1,j+1)$$

$$\gamma : H(i,j+nblig/2) = I(i,j) - I(i,j+1) + I(i+1,j) - I(i+1,j+1)$$

$$\epsilon : H(i+nblig/2,j+nblig/2) = I(i,j) - I(i,j+1) - I(i+1,j) + I(i+1,j+1)$$

avec $I(i,j)$ la valeur du signal en ligne i et colonne j, et $H(i,j)$ valeur obtenue par transformée de HAAR.

IV.2 Filtrage

- image SPOT panchromatique résolution 10 mètres rééchantillonnée en bilinéaire à 5 mètres,
- photo-aérienne scannée rééchantillonnée à 5 mètres
- image RAMSES résolution 5 mètres.

Après avoir décomposé les 3 images suivant le procédé décrit plus haut, il reste à combiner les 3 transformées. L'idée de base consiste à choisir pour chaque pixel et pour chaque niveau de décomposition, la transformée, ou la combinaison de transformées la plus appropriée. Après plusieurs essais, le filtre retenu est un ensemble de combinaisons linéaires issues de l'analyse en composantes principales (une analyse pour chaque niveau). La Photo 4 illustre l'image obtenue à partir de 9 combinaisons linaires effectuées aux 9 niveaux de transformation.

V- Classification

V.1 Constitution d'une référence

Les méthodes de classification supervisées nécessitent une référence afin d'initialiser correctement le processus et de mesurer précisément la fiabilité des résultats obtenus. La vérité terrain au moment des prises de vues est très rarement disponible. C'est pourquoi, nous avons choisi de constituer un modèle en actualisant une carte IGN au 1/25 000 à l'aide des informations dont nous disposons (photo-aériennes). Ce choix a pour conséquence fâcheuse de limiter le nombre de classes

que l'on peut définir avec certitude ; en pratique, il sera difficile de dépasser 4 ou 5 classes dans la cartographie au 1/25 000. La Photo 5 illustre cette référence.

- forêt : noir
- champ : gris clair
- eau : gris foncé
- bâtiment : blanc

V.2 Classification décentralisée "par maximum a posteriori"

V.2.1 Position du problème

Étant donnée un vecteur de mesure m comportant p variables caractérisant un individu, ici une région issue de la segmentation, il s'agit de décider de son appartenance à l'une des 4 classes répertoriées a priori, représentative de la nature du sol, ici champ (H_0), eau (H_1), forêt (H_2) et bâtiment (H_3). Le critère de décision retenue est celui du maximum a posteriori. Le théorème de Bayes permet de calculer la probabilité a posteriori d'une hypothèse, connaissant :

- les probabilités a priori de chacune des hypothèses, données par leurs fréquences relatives d'apparition au sein de l'échantillon,
- les probabilités conditionnelles $p(m/H_j)$ estimées à partir des histogrammes normalisés calculés sur une population pour laquelle on sait l'hypothèse vérifiée.

$$\forall i=0..3 \quad p(H_i/m) = \frac{p(m/H_i)p(H_i)}{\sum_{j=0}^3 p(m/H_j)p(H_j)}$$

V.2.2 Modélisation des distributions

Pour chacune des populations d'apprentissage Spot et radar, les histogrammes de 10 variables sont calculés pour toutes les

classes avec un pas de discrétisation de $\frac{\sigma}{10}$ où σ désigne l'écart-type de la distribution. Ils font apparaître des distributions unimodales. Les matrices de corrélation des 10 variables calculées sur les distributions d'apprentissage montrent que l'hypothèse d'indépendance n'est pas légitime. Le cas multivarié est très délicat à traiter sauf si l'on fait l'hypothèse gaussienne. La densité de probabilité conjointe $p(m/H)$ est alors entièrement déterminé par les moments d'ordre 1 (moyenne) et 2 (matrice de covariance).

V.2.3 Fusion bayésienne

- On pose
- D_S : décision donnée par le capteur visible
- D_R : décision donnée par le capteur radar

Dans le cas où les deux capteurs sont indépendants, ce qui peut être raisonnablement envisagé ici :

$$p(H_i / D_S, D_R) = \frac{p(D_S / H_i) p(D_R / H_i) p(H_i)}{p(D_S, D_R)}$$

Les $P(D_S/H_i)$ et $P(D_R/H_i)$ qui représentent les vraisemblances de l'hypothèse H_i dans les deux classifications peuvent être approximées par la matrice de confusion issue des classifications mono-mode.

$p(D_S, D_R)$ demeure constant quelque soit i. Il vient donc pour le maximum a posteriori :

$$\text{Décision : } \underset{i}{\text{MAX}} [p(D_S / H_i) p(D_R / H_i) p(H_i)]$$

V.2.4 Résultats

Nous présentons ici les matrices de confusion issues des classifications monomodes réalisées sur la base d'apprentissage. Le vecteur de mesure utilisé est le suivant : m = (moyenne, écart-type, homogénéité).

Figure 1 - Matrice de confusion radar

Classe	Champs	Eau	Forêt	Bâtiment
Ref Champ	65 %	7 %	24 %	2 %
Ref Eau	25 %	63 %	10 %	0 %
Ref Forêt	37 %	2 %	57 %	1 %
Ref Bâtiment	49 %	20 %	11 %	18 %

Fiabilité globale : 62 %

Figure 2 - Matrice de confusion SPOT -

classe	Champs	Eau	Forêt	Bâtiment
Ref Champ	67%	9%	19%	2%
Ref Eau	11%	76%	11%	0%
Ref Forêt	9%	3%	85%	0%
Ref Bâtiment	62%	0%	15%	21%

Fiabilité globale : 71%

Classes les mieux discriminées : 'eau' et 'forêt'.

En utilisant l'algorithme décrit ci-dessus, les résultats de la fusion bayésienne apparaissent sur la Figure 3.

Figure 3 - Matrice de confusion bimode

Classe	Champs	Eau	Forêt	Bâtiment
Ref Champ	62%	11%	18%	9%
Ref Eau	8%	85%	3%	3%
Ref Forêt	9%	6%	82%	2%
Ref Bâtiment	39%	5%	9%	48%

Fiabilité globale : 69%

Classes les mieux discriminées : 'eau', 'forêts', et 'bâtiment'.

La fusion décentralisée apparaît assez souple et didactique au sens où les points forts des deux capteurs sont distincts. De ce point de vue, la fusion bayésienne s'est révélée globalement satisfaisante.

V.3 Classification centralisée par mesure de texture

La décision centralisée est prise à partir des deux premières composantes principales issues de l'ACP des images SPOT panchromatique et SAR RAMSES.

La méthode de calcul de la distance (de type inertie) permettant l'affectation des régions à la classe la plus proche est la suivante :

Soit F'_r (resp F'_c) la matrice des cooccurrences de la région r (resp de la classe c), ordonnée par fréquences croissantes.

$$D_{KS}(r,c) = \text{Max}_{i=1}^i \text{Max}_{j=1}^j \left| \sum_{\chi=1}^{\chi=j} F'_r(t,\chi) - \sum_{\chi=1}^{\chi=j} F'_c(t,\chi) \right|$$

Figure 4 - Matrice de confusion ACP -

Classification	Champ	Eau	Forêt	Bâtiment
Ref Champ	82%	1%	10%	5%
Ref Eau	36%	60%	3%	20%
Ref Forêt	18%	1%	79%	17%
Ref Bâtiment	45%	0%	15%	38%

Fiabilité globale : 75%

Classes les mieux discriminées : 'champ'.

V.4 Conclusion

Plusieurs méthodes de classifications supervisées classiques (dont les k-plus proches voisins) ont été testées sur ce site (Figure 5). Tous types de classification confondus, les résultats de la classification 'par texture' sont les meilleurs obtenus (Photo 6). L'ACP permet d'optimiser la classification quelle que soit la classe considérée. A méthode comparable (maximum a posteriori), la fusion centralisée obtient de meilleurs résultats comparativement à une fusion décentralisée.

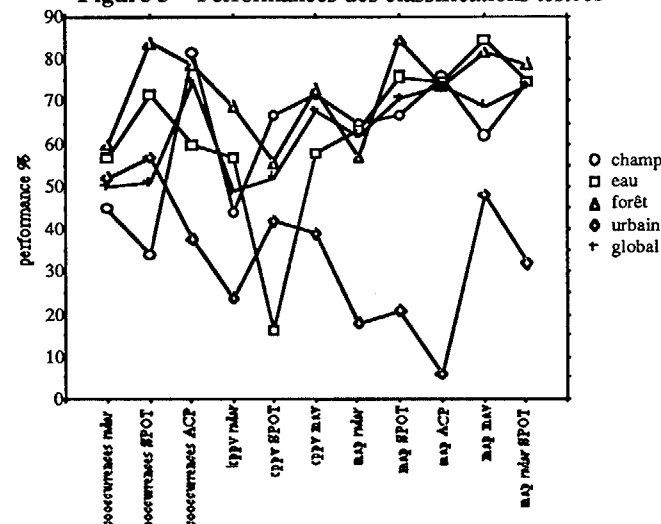
De par leur faible représentation dans l'image, les bâtiments sont les moins bien classés. La sous représentation de la classe 'eau' s'explique par le choix de la référence issue de la carte IGN et de l'image SPOT. Les classifications concernant le poste 'forêt' apparaissent par endroit plus valides que la référence si l'on regarde attentivement les deux images. Cependant n'ayant pu nous rendre sur le terrain, nous n'avons pas voulu modifier la référence (issue de la carte et de la photographie aérienne), ce qui a pour conséquence une sous-estimation du pourcentage de 'forêt' bien classé.

VI Conclusions et perspectives

Pour les classifications envisagées sous l'aspect bimode SPOT Panchromatique (10m), RAMSES (5m), l'apport du radar est indiscutable. Des informations multispectrales SPOT, très riches, changeraient sans doute la nature du problème.

La question de la prise en compte d'images multi-dates est importante et n'a pu être résolue ici. Il est en effet très délicat, de décider si une affectation différente des deux capteurs est due à un changement d'occupation du sol ou à une contradiction entre capteurs. La prise en compte de plus nombreuses sources d'information (panchromatique, multispectrale, et radar SAR multi-dates) sera certainement plus apte à lever ces ambiguïtés.

Figure 5 - Performances des classifications testées -



L'analyse en composantes principales ("fusion au niveau pixel") s'avère être un instrument adapté à la prise en compte d'informations bimodes. De toutes celles testées, la classification centralisée s'est révélée la plus satisfaisante par rapport à la référence disponible.

La méthode de fusion pyramidale bien qu'à l'heure actuelle peu convaincante, semble une alternative prometteuse face à une simple approche de type combinaison linéaire mono-échelle.

La classification 'région' est plus pertinente qu'une classification 'pixel'. Ceci n'est guère surprenant, la littérature consacrée à ce sujet en télédétection étant relativement consensuelle (en particulier en ce qui concerne les milieux complexes). De plus, si l'on parvient à mieux segmenter les images (respect de la forme du parcellaire en particulier), la prise en compte de variables morphologiques pourrait améliorer sensiblement les résultats.

Par ailleurs les techniques de l'incertain pour combiner des évaluations mono-mode peuvent être envisagées maintenant que les performances de la fusion bayésienne ont été évaluées. La prise en compte de données exogènes telles que le relief (ou un modèle d'éclairement), une ancienne carte ou des informations contextuelles pourraient avantageusement être gérées par ce type de méthode.

VII- Références bibliographiques:

- [HOU91] HOUZELLE S., GIRAUDON G., 1991 - Extraction d'objets par fusion de données d'images satellites, ORASIS St Malo, 21 P.
- [WEL88] WELCH R., ELHER M., 1988 - Cartographic feature extraction with integrated SIR-B and Landsat TM image, International Journal of Remote Sensing, Vol 9, N°5 873-889 pp.
- [YES91] YESOU H., BESNUS Y., 1991 - Intégration des données satellitaires Seasat et SPOT, contribution à la cartographie géologique structurale en milieu tempéré. Colloque Courchevel 1991.



Photo 1 : SPOT panchromatique
© CNES 1990 - Distribution SPOT IMAGE



Photo 2 : SAR RAMSES



Photo 3 : suppression des ombres

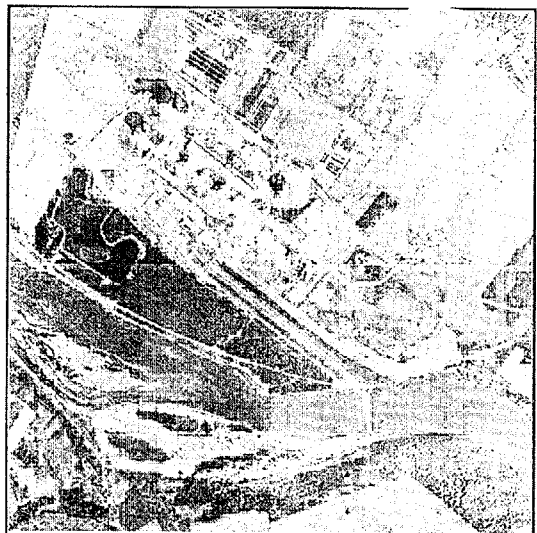


Photo 4 : transformée pyramidale

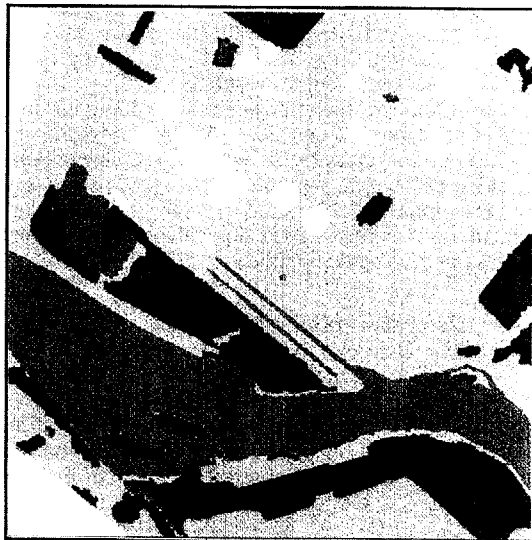


Photo 5 : référence



Photo 6 : classification par texture