





Soutenance de Thèse

Méthodes d'analyse de texture pour la cartographie d'occupations du sol par télédétection très haute résolution

Application à la forêt, la vigne et les parcs ostréicoles

OLIVIER REGNIERS

Thèse réalisée sous la direction de :

Christian Germain (IMS, Bordeaux Sciences Agro) Lionel Bombrun (IMS, Bordeaux Sciences Agro)

11 décembre 2014

Contexte – Image de télédétection optique THR (panchromatique)



Contexte – Image de télédétection optique THR

Exemples d'occupations du sol texturées



Contexte – Analyse texturale



3

Méthode de référence : Matrice de co-occurrence (GLCM)

GLCM = Nombre d'occurrences de toutes les paires de valeurs de niveaux de gris situées à une distance d dans l'image dans une direction dx,dy



90°

45°

→ 0°

135°

Transformée en ondelettes

Transformée en ondelettes discrète (DWT)

- \rightarrow Combinaison de filtres passe-haut (H) et passe-bas (L)
- ightarrow Sous-bandes d'ondelettes orthogonales et indépendantes
- → Représentation multi-échelles et multi-orientations du contenu fréquentiel
- → Limitée en termes de sélectivité directionnelle (3 orientations) et d'invariance à la translation



LL ₂	HL_2	ы
LH ₂	HH_2	
LH1		ΗH1



Transformée en ondelettes

Transformée en ondelettes discrète (DWT)

- \rightarrow Combinaison de filtres passe-haut (H) et passe-bas (L)
- → Sous-bandes d'ondelettes orthogonales et indépendantes
- → Représentation multi-échelles et multi-orientations du contenu fréquentiel
- → Limitée en termes de sélectivité directionnelle (3 orientations) et d'invariance à la translation







Ondelette dyadique → sous-échantillonnage d'un facteur 2 à chaque échelle Ondelette-mère = Daubechies db4

Autres transformées en ondelettes : ondelettes stationnaires, pyramides orientées, etc.









Modèle Gaussien multivarié

$$\boldsymbol{k} \sim \mathcal{N}(\boldsymbol{0}, \mathbf{M}) \qquad \left[\hat{\boldsymbol{M}} \right]_{SCM} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{k}_{i} \boldsymbol{k}_{i}^{'} \qquad \boldsymbol{\rightarrow} \mathbf{SCM}$$

Modèles SIRV (Spherically Invariant Random Vectors)

$$\boldsymbol{k} = \sqrt{\tau} \boldsymbol{z} \qquad \boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \mathbf{M})$$

Si loi Gamma Inverse sur τ, vecteur SIRV **k** suit loi G0

Si Dirac sur τ, vecteur SIRV k suit loi Gaussienne

$$\left[\hat{\boldsymbol{M}}\right]_{FP} = \frac{p}{N} \sum_{i=1}^{N} \frac{\boldsymbol{k}_{i} \boldsymbol{k}_{i}}{\boldsymbol{k}_{i} \left[\hat{\boldsymbol{M}}\right]_{FP}^{-1} \boldsymbol{k}_{i}} \qquad \hat{\boldsymbol{\tau}}_{i} = \frac{\boldsymbol{k}_{i} \left[\boldsymbol{M}\right]^{-1} \boldsymbol{k}_{i}}{p}$$



Modèles sur copules

$$F(\mathbf{k}_{i}) = F(k_{i,1}, \dots, k_{i,p}) = C(F_{i,1}(k_{i,1}), \dots, F_{i,p}(k_{i,p}))$$

 $\left[\hat{\boldsymbol{M}}\right] = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{k}_{i}^{G} \boldsymbol{k}_{i}^{G'}$

Gamma multivariée sur copule Gaussienne → GCG Estimateur de la matrice de covariance (+ Estimateur autres paramètres) → SIGNATURE TEXTURALE

Indexation d'images (CBIR)



Notion de similarité

GLCM → distance de Mahalanobis

SCM, SIRVgauss, SIRVg0, GCG

 \rightarrow divergence de Kullback-Leibler

→ distance géodésique

Indépendance des sous-bandes

$$D(I_1 || I_2) = \sum_{sc=1}^{Nsc} \sum_{or=1}^{Nor} D(W_{1sc,or} || W_{2sc,or})$$

Taux de Reconnaissance Moyen (ARR)

$$ARR = \frac{1}{N_T} \sum_{q=1}^{N_T} \frac{n_q(N_C)}{N_C}$$

 N_T = Nombre total d'images

 N_c = Nombre d'images par classe

 $n_q(N_C)$ = Nombre d'images appartenant à la même classe C que l'image requête parmi les N_c les plus proches

Courbe Rappel/Précision

Rappel =
$$R(k) = \frac{1}{N_T} \sum_{q=1}^{N_T} \frac{n_q(k)}{N_C}$$

$$Pr\acute{e}cision = P(k) = \frac{1}{N_T} \sum_{q=1}^{N_T} \frac{n_q(k)}{k}$$

Similaire à la courbe **ROC** (TVP vs TFN) **Mais** Plus informative si base de données déséquilibrée



Stratégies d'évaluation des performances de classification



k plus proches voisins (k-NN), maximum de vraisemblance (ML), machine à vecteurs de support (SVM)

Thématique Forêt – Contexte



Forêt des Landes de Gascogne

- 10⁶ hectares
- forêt mono-spécifique équienne
- 90% de pins maritimes cultivés
- organisation parcellaire morcelée
- séquelles de la tempête Klaus (2009) et de la tempête Martin (1999)



Thématique Forêt – Contexte



d'âges de peuplements forestiers (forêt mono-spécifique de pins maritimes)

Thématique Forêt – Données Image



3 images Pléiades (RTU programme ORFEO – CNES)

1 image GeoEYE-1 (Telespazio – EarthLab)

 \rightarrow Résolution spatiale PAN = 0,5 m

	Pléiades 24/06/12	Pléiades 08/08/12	Pléiades 20/02/13	GeoEYE-1 03/04/13
Angle d'incidence	27,5°	15°	22°	20,5°
GSD PAN	0,83 m	0,74 m	0,78 m	0,45 m
Elévation solaire	66°	59,5°	32,5°	49°

Thématique Forêt – Base de données de patchs de texture

Classe d'âge 1 Classe d'âge 2 Classe d'âge 3 Coupe-Rase **ATTENTION** : âge peuplements = variable continue

Base de données réalisée par photo-interprétation (dimension patchs = 256 x 256 pixels) BD

Sensibilité à l'anisotropie



Seconde base de données BDrot

Evaluation des performances des modèles en indexation

Estimation de l'orientation par un tenseur de structure

GLCM

Descripteurs moyennés sur 4 orientations (0°, 45°, 90°, 135°)

 \rightarrow Invariance à la rotation

Sélection des descripteurs par ACP

 \rightarrow homogénéité, entropie, corrélation, moyenne

Sélection de la distance entre paires de pixels

 \rightarrow Distance de 2 pixels plus performante



Modèles multivariés

Transformée en ondelettes discrète (DWT)

2 échelles x 3 orientations = 6 sous-bandes

Dimension sous-bandes échelle 1 = 128 x 128, échelle 2 = 64 x 64

Voisinage de dépendance spatiale = 3 x 3



Thématique Forêt – Résultats d'indexation



Thématique Forêt – Résultats d'indexation



19

Thématique Forêt – Base de données de référence





Contours de 179 peuplements (coupe-rases incluses) \rightarrow 0 à 50 ans par tranche d'âge de 5 ans

Classe d'âge 1 : peuplements de 0 à 9 ans Classe d'âge 2 : peuplements de 10 à 19 ans Classe d'âge 3 : peuplements de plus de 19 ans

Thématique Forêt – Modalités de caractérisation des peuplements



SP+COL : Dépendance spatiale **SP** + Dépendance couleur **COL**



Comparaison de deux classifieurs :

k-NN vs SVM



Thématique Forêt – Résultats de classification k-NN (k = 5)



Meilleures performances avec modèle SIRVgauss

Meilleures performances des modèles multivariés (SP)

- \rightarrow Seuil de distance faible (k = 5)
- → Textures des peuplements hétérogènes

Producer's Accuracy par classe – modèle SIRVgauss

	Classe	Classe	Classe	Coupe-	04
	d'âge 1	d'âge 2	d'âge 3	Rase	UA
SP	73.24	63.86	94.55	56.62	73.47
COL	66.06	73.69	81.3	32.5	68.89
SP+COL	72.47	79.66	86	47.87	75.57
SProt	83.09	69	82.1	47.37	75.24
SProt+COL	86.91	79.62	85.35	50.75	81.07

OA = Overall Accuracy

Effet synergétique de la dépendance couleur et de la rotation (SProt+COL)

→ Caractéristiques propres à chacune des classes

Meilleure capacité des modèles multivariés (SIRVgauss) à capturer des motifs texturaux complexes

Thématique Forêt – Résultats de classification SVM



Noyau Gaussien
$$K = \exp\left(\frac{-(distance)^2}{\sigma^2}\right)$$

Modèle SIRVgauss toujours plus performant

Meilleure discrimination des classes anisotropes

Meilleure valorisation des paramètres de la dépendance couleur

Producer's Accuracy par classe – modèle SIRVgauss

	Classe d'âge 1	Classe d'âge 2	Classe d'âge 3	Coupe- Rase	OA
SP	91.68	78.76	85	40	81.55
COL	94.53	75	71.1	7.37	75.5
SP+COL	93.47	84.65	84.25	39.37	83.89
SProt	91.91	79.17	67.15	34.62	77.37
SProt+COL	94.18	84.1	84.4	39.62	84.02

OA = Overall Accuracy

Confusion liée à l'anisotropie réduite + valorisation de la dépendance couleur par l'utilisation des SVMs

Performances de classification plafonnée à <u>+</u> 85%

→ Présence significative de peuplements d'aspect aberrant



Limites de l'a priori principal -> relation entre âge des peuplements et distribution spatiale des arbres

CONCLUSIONS

- Modèles multivariés (SIRVgauss) plus performants que GLCM → textures hétérogènes
- Caractéristiques propres à chacune des classes \rightarrow orientation, couleur, texture
- Classifieur SVM plus efficace que k-NN
- Hypothèse *a priori* pas aussi simple

PERSPECTIVES

- Modélisation par classe plutôt que modélisation globale
 - → Tests de Goodness-of-Fit
 → Nombre d'échelles, types de modèle différents
- Autres variables de structure forestière