

Modèles probabilistes Riemanniens de tenseurs de structure pour le traitement d'images texturées.

Roxana ROSU¹, Lionel BOMBRUN^{2,3}, Salem SAID², Olivier REGNIERS², Marc DONIAS^{2,4}, Jean-Pierre DA COSTA^{2,3}

¹Univ. Bordeaux, IMS, UMR 5218, F-33400 Talence, France

²CNRS, IMS, UMR 5218, F-33400 Talence, France

³Bordeaux Sciences Agro, F-33175 Gradignan, France

⁴Bordeaux INP, IMS, F-33400 Talence, France

prenom.nom@ims-bordeaux.fr

Résumé – Nous proposons dans cet article de nouveaux outils mathématiques pour le traitement d'images texturées anisotropes. Nous introduisons une nouvelle classe de modèles probabilistes de mélanges basés sur une loi dite *gaussienne* dans la variété Riemannienne des matrices symétriques définies positives. Nous montrons la pertinence de ces outils pour la modélisation de tenseurs de structure locaux calculés sur des textures réelles et illustrons leurs potentialités dans un contexte de classification d'images de couverts forestiers en télédétection haute-résolution.

Abstract – New tools are introduced for the handling of anisotropic textures in image analysis. In particular, we present a new class of mixture models derived from a *Gaussian* distribution within the Riemannian manifold of symmetric positive definite matrices. These models are proved to efficiently model the local structure tensors computed on real textures. Their potentialities are illustrated in the context of high resolution remote sensed forest image classification.

1 Introduction

L'anisotropie est une des principales caractéristiques utilisées en analyse d'image texturées. Justifiée par des travaux fondateurs sur la perception des textures, son utilisation est également très courante en vision par ordinateur. Les approches proposées pour capturer et décrire l'anisotropie des textures sont très diverses, se basant par exemple sur les statistiques spatiales de niveaux de gris [3], sur l'analyse du spectre par des filtres de Gabor [9] ou par la décomposition de Wold [8].

Un autre outil standard pour l'étude de l'anisotropie est le tenseur de structure [2]. Résumant l'information relative à la distribution des gradients au voisinage d'un pixel, il permet d'extraire une direction dominante et d'évaluer la cohérence des gradients autour de cette direction. S'il s'agit d'un outil courant en analyse de textures, assez peu d'approches s'appuient sur sa modélisation probabiliste. Dans un contexte d'estimation-segmentation jointes de champs d'orientation, Da Costa et al. [4] modélisent les orientations des tenseurs comme des données scalaires définies sur le cercle. Dans un contexte de segmentation, Rousson et al. [11] utilisent une représentation vectorielle du tenseur de structure et modélisent les tenseurs d'une région selon une loi normale multi-variée tandis que de Luis-García et al. [5] exploitent une mesure intrinsèque de distance entre tenseurs dont ils modélisent la dispersion par une loi normale. L'utilisation de tenseurs de structure – en tant que matrices symétriques définies positives (SDP) – pose un problème formel de géométrie de l'information et ne peut se faire avec les outils classiques de géométrie euclidienne. L'adoption d'outils probabilistes adéquats [1, 7, 6] doit tenir compte du fait que ces matrices appartiennent à une variété Riemannienne.

Dans cet article, nous proposons de nouveaux outils pour la description probabiliste de matrices SDP et les appliquons aux tenseurs de structure. A l'instar de [7, 5], ces outils s'appuient sur la distance de Rao dont l'invariance affine nous permet d'introduire un modèle dit *gaussien* de matrices SDP et différents outils dérivés (estimateurs, modèles de mélange et distances). En section 2, nous présentons synthétiquement ces outils et les appliquons en section 3 à la modélisation de tenseurs de structures. Une évaluation expérimentale de l'approche est réalisée dans un contexte de classification d'images de télédétection.

2 Modèles statistiques de matrice SDP

2.1 L'espace \mathcal{P}_2

Considérons l'espace \mathcal{P}_2 des matrices réelles 2×2 symétriques définies positives. \mathcal{P}_2 est une variété de dimension 3 dont chaque représentant $Y = \begin{pmatrix} y_{11} & y_{12} \\ y_{12} & y_{22} \end{pmatrix}$ peut être paramétré par un triplet (r_1, r_2, θ) obtenu par décomposition spectrale :

$$Y(r_1, r_2, \theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} e^{r_1} & 0 \\ 0 & e^{r_2} \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (1)$$

avec $(r_1, r_2) \in \mathbb{R}$, $r_1 \geq r_2$ et $\theta \in [0, 2\pi[$.

Un point intéressant concernant l'espace \mathcal{P}_2 , est qu'il est homogène sous l'action de GL_2 , groupe des matrices réelles 2×2 non singulières : $\forall (g, Y) \in GL_2 \times \mathcal{P}_2$, $g^\dagger Y g \in \mathcal{P}_2$, où g^\dagger est la transposée de g . Il ressort que pour tout $Y_1, Y_2 \in \mathcal{P}_2$, il existe $g \in GL_2$ telle que $Y_2 = g^\dagger Y_1 g$. Dans ce qui suit, nous nous intéresserons en particulier au cas $Y_1 = Y$ et $Y_2 = Id$ (matrice identité) qui, pour $g = Y^{-\frac{1}{2}}$, vérifient $Id = g^\dagger Y g$.

2.2 Une distribution gaussienne sur \mathcal{P}_2

Introduisons la loi $G(\bar{Y}, \gamma)$ dite *gaussienne* sur \mathcal{P}_2 définie par sa densité de probabilité :

$$p(Y|\bar{Y}, \gamma) = Z^{-1}(\bar{Y}, \gamma) e^{-\frac{d^2(Y, \bar{Y})}{2\gamma^2}} \quad (2)$$

où $\bar{Y} \in \mathcal{P}_2$ et $\gamma \in \mathbb{R}$ représentent un barycentre et une dispersion et $Z(\bar{Y}, \gamma)$ est une constante de normalisation. La fonction $d : \mathcal{P}_2 \times \mathcal{P}_2 \rightarrow \mathbb{R}_+$ désigne la distance de Rao :

$$d^2(Y_1, Y_2) = \text{Tr} \left[\log^2 \left(Y_1^{-\frac{1}{2}} Y_2 Y_1^{-\frac{1}{2}} \right) \right]. \quad (3)$$

On peut montrer que d est invariante sous l'action de GL_2 . Il découle que $Z(\bar{Y}, \gamma)$ ne dépend pas du barycentre \bar{Y} et peut s'écrire $Z(\gamma) = 4\pi^2 \gamma^2 e^{\gamma^2/4} \text{erf}(\gamma/2)$, où erf est la fonction d'erreur. Une autre conséquence est que pour tout $Y \sim G(\bar{Y}, \gamma)$, $p(Y|\bar{Y}, \gamma) = p(Y^*|Id, \gamma)$ avec $Y^* = \bar{Y}^{-\frac{1}{2}} Y \bar{Y}^{-\frac{1}{2}}$.

Nous appellerons *distribution centrée* la distribution $G(Id, \gamma)$. En introduisant les coordonnées polaires $t = r_1 + r_2$, $r = r_1 - r_2$ et θ , il est possible de montrer que :

$$p(Y|Id, \gamma) = p(r, t, \theta) \propto e^{-\frac{r^2}{4\gamma^2}} \sinh\left(\frac{r}{2}\right) \times e^{-\frac{t^2}{4\gamma^2}} \times \frac{1}{2\pi}. \quad (4)$$

Il apparaît ainsi que r , t et θ sont indépendants, que t est de distribution normale et que θ est de distribution uniforme.

Considérons un échantillon Y_1, \dots, Y_N issu d'une loi $G(\bar{Y}, \gamma)$. D'après (2), la log-vraisemblance de l'échantillon s'écrit :

$$\mathcal{L}(\bar{Y}, \gamma) = -N \log Z(\gamma) - \frac{1}{2\gamma^2} \sum_{i=1}^N d^2(Y_i, \bar{Y}) \quad (5)$$

qui conduit à une estimation de \bar{Y} au sens du maximum de vraisemblance (MV) :

$$\hat{\bar{Y}} = \underset{Y \in \mathcal{P}_2}{\text{argmin}} \sum_{i=1}^N d^2(Y_i, Y) \quad (6)$$

qui n'est autre que le barycentre des Y_i au sens de la distance de Rao obtenu à l'aide d'un algorithme Riemannien de descente de gradient. Enfin, une estimation $\hat{\gamma}$ de la dispersion au sens du MV s'obtient en résolvant l'équation non linéaire d'inconnue γ :

$$\frac{1}{N} \sum_{i=1}^N d^2(Y_i, \bar{Y}) = \gamma^3 \frac{d}{d\gamma} \log Z(\gamma). \quad (7)$$

2.3 Mélange de lois gaussiennes sur \mathcal{P}_2

Il est possible de définir un modèle plus général de loi sous la forme d'un mélange de K lois gaussiennes $p(Y|\bar{Y}_k, \gamma_k)$:

$$p(Y|(\omega_k, \bar{Y}_k, \gamma_k)_{k=1, \dots, K}) = \sum_{k=1}^K \omega_k p(Y|\bar{Y}_k, \gamma_k) \quad (8)$$

où les ω_k sont des poids positifs de somme égale à 1. L'estimation des paramètres du mélange peut être conduite à l'aide d'un algorithme de type *Expectation-Maximization*.

2.4 Distances entre lois sur \mathcal{P}_2

Considérons deux échantillons issus de deux lois p et q inconnues. Une option pour les comparer consiste à calculer une

distance entre les lois \hat{p} et \hat{q} estimées à partir de ces mêmes échantillons. Dans le cas de matrices SPD, il s'agirait par exemple d'estimer les paramètres des gaussiennes ou des mélanges de gaussiennes sur \mathcal{P}_2 . Une distance possible est la version symétrique de la divergence de Kullback-Leibler :

$$D_{KLS}(\hat{p}, \hat{q}) = D_{KL}(\hat{p}, \hat{q}) + D_{KL}(\hat{q}, \hat{p}) \quad (9)$$

$$\text{où } D_{KL}(\hat{p}, \hat{q}) = \int_x p(x) \log \frac{p(x)}{q(x)} dx \quad (10)$$

En l'absence de forme analytique pour ces divergences, une estimation est possible à condition de simuler des échantillons :

$$\hat{D}_{KLS}(\hat{p}, \hat{q}) = \frac{1}{M} \sum_{m=1}^M \log \frac{\hat{p}(x_m)}{\hat{q}(x_m)} + \frac{1}{M} \sum_{m=1}^M \log \frac{\hat{q}(y_m)}{\hat{p}(y_m)} \quad (11)$$

où les x_m et y_m sont tirés selon les lois p et q . La génération d'un échantillon selon une loi $G(\bar{Y}, \gamma)$ s'appuie sur les propriétés énoncées à la section 2.2. Le principe est de générer x_1^*, \dots, x_M^* selon $G(Id, \gamma)$ puis de les translater autour de \bar{Y} selon la procédure : $x_m = \bar{Y}^{1/2} x_m^* \bar{Y}^{1/2}$. Ce principe peut aisément être étendu au cas d'un mélange de gaussiennes.

3 Modélisation de tenseurs de structure

3.1 Estimation du tenseur de structure

Le tenseur de structure Y d'une image I est communément défini comme la covariance locale du gradient ∇I [2, 5] :

$$\nabla I = [I_x, I_y]^\dagger = [I * G_x, I * G_y]^\dagger \quad (12)$$

où $*$ est l'opérateur de convolution et G_x et G_y sont des noyaux dérivatifs gaussiens [38,39] de paramètre σ_G . Le tenseur de structure est estimé en lissant le produit $\nabla I \nabla I^\dagger$ par un noyau convolutif gaussien W_T de paramètre σ_T :

$$Y = W_T * \nabla I \nabla I^\dagger = W_T * \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix} \quad (13)$$

Le lissage spatial confère une certaine robustesse au bruit et permet de distribuer l'information fournie par le gradient de part et d'autre des contours, y compris près des extrema locaux de l'image où le gradient est nul. Il permet aussi de paramétrer l'échelle à laquelle l'orientation est estimée. Dans la suite, σ_T sera choisi en fonction de la taille des motifs texturaux présents dans l'image.

L'exploitation du tenseur de structure s'appuie généralement sur sa diagonalisation. Les vecteurs propres unitaires u_1 et u_2 associés aux valeurs propres $\lambda_1 \geq \lambda_2$ indiquent, en tout point de l'image, les orientations orthogonale et tangente aux motifs. La présence d'un motif anisotrope se traduit par des valeurs propres $\lambda_1 \gg \lambda_2$ très différentes alors que des valeurs propres semblables $\lambda_1 \simeq \lambda_2$ indiquent l'absence de tendance directionnelle locale. Dans certains situations dégénérées, le tenseur n'est pas strictement défini positif (e.g. λ_2 nul), auquel cas les développements vus en section 2 ne s'appliquent pas.

Notons que les tenseurs calculés en des pixels voisins ne sont pas indépendants du fait du filtrage introduit par W_T . D'un

point de vue formel, l'utilisation des estimées \hat{Y} et $\hat{\gamma}$ est donc abusive. Une alternative est d'extraire un sous-ensemble de tenseurs calculés en des points suffisamment distants pour être considérés indépendants. En pratique, cette opération pose cependant des problèmes d'estimation car elle réduit considérablement la taille des échantillons utilisés.

3.2 Expérimentation sur des images texturées

Nous avons conduit des expérimentations sur une base de données de télédétection contenant des images de forêts de pin maritime¹. Ces images sont réparties en trois classes d'âge : C_1 (0–9 ans), C_2 (10–19 ans), C_3 (>20 ans) et une dernière classe C_4 de coupes rases. Dans la figure 1, on montre pour chaque classe un exemple d'image et de champ d'orientation associé. Les orientations sont issues des tenseurs de structure, calculés avec $\sigma_G = 1.9$ et $\sigma_T = 3$ ². Les textures de la première classe sont anisotropes. On y distingue clairement les rangs d'arbres. Du fait du grossissement des houppiers, les textures de C_2 présentent une anisotropie plus faible tandis que celles de C_3 et C_4 perdent leur direction privilégiée comme en témoigne la variabilité observée sur les champs d'orientation de la figure 1.

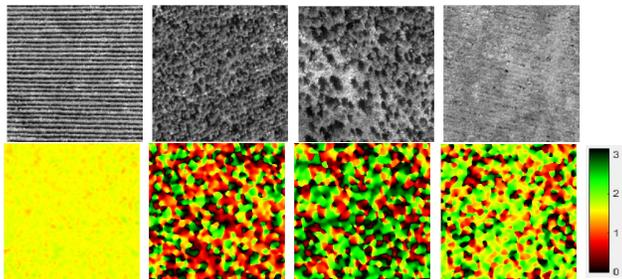


FIGURE 1 – Textures originales (en haut) pour les classes C_1 à C_4 (de gauche à droite) et leurs champs d'orientation (en bas).

Le champ de tenseurs de structure est modélisé par une loi gaussienne dont on estime le barycentre \bar{Y} et la dispersion γ selon (6) et (7). Après centrage des tenseurs, on calcule leurs paramètres r , t et θ dont on compare les distributions avec celles de données simulées selon $G(\bar{Y}, \gamma)$ (cf. figure 3). On constate que les distributions expérimentales de r et t sont très proches de celles simulées. En revanche, la distribution de θ est très différente de la distribution uniforme attendue pour $G(\bar{Y}, \gamma)$ et observée pour les données simulées. Ceci nous amène à modéliser le champ tensoriel par un mélange de lois gaussiennes.

Dans la figure 2, pour une texture de chaque classe, nous traçons la variation, en fonction de K , du logarithme du rapport des vraisemblances d'un mélange de K gaussiennes et d'une gaussienne seule. La courbe croissante montre que le mélange de lois gaussiennes représente plus fidèlement la distribution des tenseurs de structures expérimentaux. C'est également ce

que l'on observe à la figure 3, où il apparaît que les données simulées par un mélange de gaussiennes sont plus proches des tenseurs expérimentaux.

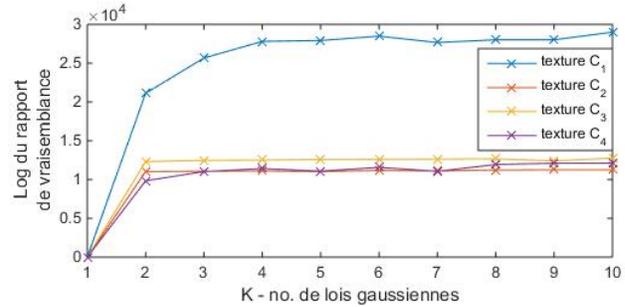


FIGURE 2 – Variation, en fonction de K , du logarithme du rapport de vraisemblance entre un mélange de K lois gaussiennes et une loi gaussienne seule.

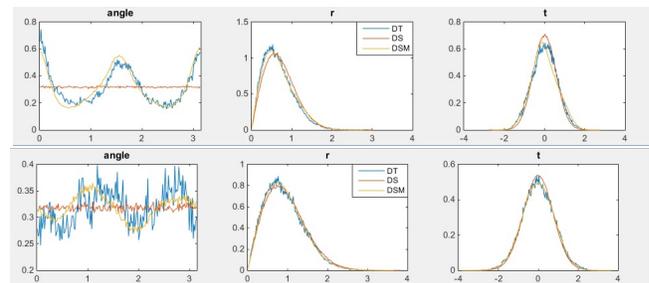


FIGURE 3 – Distribution des paramètres du tenseur centré pour les textures 1 (en haut) et 2 (en bas) : DT données texture, DS - données simulées pour une gaussienne et DSM données simulées pour un mélange de 5 gaussiennes

3.3 Classification de textures

Afin de démontrer les potentialités des modèles, nous proposons de mener une expérimentation sur l'ensemble de la base de données. Le principe retenu est celui d'une recherche d'images par le contenu. Dans une première étape, on extrait une signature pour chaque image de la base. Puis, on considère successivement chacune des images comme une image requête et, en utilisant une mesure de similarité, on recherche les images les plus similaires dans la base. Une présentation détaillée du protocole est disponible dans [10].

Dans notre cas, la signature est donnée par le barycentre et la dispersion des tenseurs de l'image, modélisés par une loi gaussienne $G(\bar{Y}, \gamma)$. L'orientation des rangs de forêt ne pouvant pas être un facteur de discrimination, cette signature doit être invariante en rotation. Cette invariance est assurée en substituant au barycentre \bar{Y} la matrice diagonale constituée de ses valeurs propres $\bar{\Lambda} = \begin{pmatrix} \bar{\lambda}_1 & 0 \\ 0 & \bar{\lambda}_2 \end{pmatrix}$, où $\bar{\lambda}_1 \geq \bar{\lambda}_2$. Ceci revient à appliquer la même rotation à tous les tenseurs d'une image de sorte qu'ils aient, en moyenne, une orientation verticale. Afin de mesurer la similarité entre deux images de la base, la distance D_{KLS} est utilisée comme évoqué à la section 2.4.

Les performances des modèles sont évaluées à l'aide du taux

1. Les auteurs remercient le CNES et l'équipe de la Recette Thématique Utilisateurs pour l'accès aux images PLEIADES.

2. Les échelles des noyaux gaussiens ont été choisies conformément à la taille des structures présentes dans les images et correspondent à un optimum des résultats de classification.

de reconnaissance moyen (TRM) défini par :

$$TRM = \frac{1}{N_I} \sum_{i=1}^{N_I} \frac{n_i(N_C)}{N_C} \quad (14)$$

où N_I est le nombre total d'images dans la base de données, N_C est le nombre d'images par classe et $n_i(N_C)$ est le nombre d'images qui appartiennent à la même classe que l'image requête i parmi les N_C images les plus proches.

La base de données contient 470 d'images, dont 140 appartient à C_1 , 109 à C_2 , 181 à C_3 et 40 à C_4 . En raison de la grande disparité du nombre d'images de chaque classe, afin de ne pas biaiser les résultats d'indexation, 25 images sont tirées aléatoirement dans chaque classe et le taux de reconnaissance est calculé pour la base de 100 images ainsi construite. L'opération est répétée 100 fois et les résultats sont moyennés.

Le tableau 1 montre les résultats obtenus en modélisant les tenseurs par une seule gaussienne (G) ou par des mélanges de 3 ou 5 lois (resp. 3G et 5G). Les distances d_{KLS} sont estimées en générant $N = 1000$ échantillons. Comme l'on pouvait s'y attendre, nous constatons une légère augmentation du taux de reconnaissance moyen lorsqu'on considère un mélange de gaussiennes. C'est le cas, en particulier, pour la classe 4.

TABLE 1 – TRM (%) pour toutes les classes (TC) et pour chaque classe d'âge : C_1 , C_2 , C_3 et C_4

	TC	C_1	C_2	C_3	C_4
G	65.3 ± 2	44.3 ± 4	56.8 ± 4	81.6 ± 4	78.9 ± 3
3G	67.1 ± 2	44.4 ± 3	57.3 ± 5	84.1 ± 4	82.4 ± 4
5G	67.4 ± 2	45.5 ± 4	57.4 ± 4	83.3 ± 4	83.3 ± 4
Cooc₁	71.8 ± 3	53.6 ± 5	75.5 ± 5	78.4 ± 5	78.9 ± 4
Cooc₂	65.7 ± 2	40 ± 3	65.4 ± 5	78.3 ± 5	79.5 ± 4

A titre de comparaison, nous présentons les résultats obtenus en considérant comme signature un jeu de quatre indicateurs issus des matrices de co-occurrence (Cooc₁). Basée sur de simples statistiques locales de niveaux de gris, cette approche s'est révélée la plus performante face notamment à certaines méthodes fondées sur la modélisation de coefficients d'ondelettes [10]. Ces bonnes performances se confirment ici, le taux de reconnaissance dépassant celui de l'approche tensorielle pour les classes C_1 et C_2 en particulier. Notons toutefois que le jeu d'indicateurs utilisé dans [10] comprend notamment la moyenne des niveaux de gris, rendant l'approche sensible aux variations de luminosité. Les résultats présentés en dernière ligne (Cooc₂) montrent qu'un simple centrage des images conduit à une chute des performances, ce qui n'est pas le cas pour le tenseur de structure, insensible à la luminance moyenne.

4 Conclusion

Nous avons présenté une nouvelle approche pour la modélisation probabiliste du tenseur de structure, outil couramment utilisé en analyse d'images pour la caractérisation de textures anisotropes. Cette approche s'appuie sur la définition d'une loi

dite *gaussienne* sur la variété Riemannienne des matrices symétriques définies positives. Nous avons en outre introduit un modèle de mélange ainsi qu'une méthode d'estimation, une mesure de distance et une procédure de simulation. La capacité de ces modèles à représenter fidèlement des tenseurs de structure a été démontrée sur un jeu de données réelles constitué d'images de télédétection forestière, au travers notamment d'un protocole de classification. Les principales perspectives de ce travail concernent la confrontation à des approches probabilistes alternatives et la généralisation des modèles à des matrices de dimension supérieure qui permettront de traiter le cas d'images volumiques ou celui de tenseurs étendus.

Références

- [1] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Log-euclidean metrics for fast and simple calculus on diffusion tensors. *Magn. Reson. Med.*, 56(2) :411–421, 2006.
- [2] J. Bigün, G.H. Granlund, and J. Wiklund. Multi-dimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(8) :775–789, 1991.
- [3] D. Chetverikov. Texture analysis using feature-based interaction maps. *Pattern Recognition*, 32 :487–502, 1999.
- [4] J.-P. Da Costa, F. Galland, A. Roueff, and C. Germain. Unsupervised segmentation based on von mises circular distributions for orientation estimation in textured images. *Journal of Electronic Imaging*, 21(2) :1–7, 2012.
- [5] R. de Luis-García, R. Deriche, and C. Alberola-López. Texture and color segmentation based on the combined use of the structure tensor and the image components. *Signal Processing*, 88 :776–795, 2008.
- [6] I. L. Dryden, A. Koloydenko, and D. Zhou. Non-euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *The Annals of Applied Statistics*, 3(3) :1102–1123, 2009.
- [7] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the manifold of multivariate normal distributions : Theory and application to diffusion tensor mri processing. *J. Math. Imaging Vis.*, 25(3) :423–444, 2006.
- [8] F. Liu and R.W. Picard. A spectral 2-d wold decomposition algorithm for homogeneous random fields. In *Proc. of ICASSP*, volume 6, pages 3501–3504 vol.6, 1999.
- [9] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(8) :837–842, 1996.
- [10] O. Regniers. *Méthodes d'analyse de texture pour la cartographie d'occupations du sol par télédétection THR - Application à la forêt, la vigne et les parcs ostréicoles*. PhD thesis, Université de Bordeaux, 2014.
- [11] M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *Proc. of CVPR*, volume 2, pages 699–704, 2003.