

1. Introduction.

La segmentation des textures couleur est un problème complexe. De nombreuses applications nécessitent l'utilisation de méthodes de segmentation précises qui soient également stables sous différentes contraintes. De telles applications utilisent généralement des méthodes de segmentation supervisées [1,2]. En segmentation supervisée, un échantillon connu de texture couleur est utilisé pour identifier et/ou extraire les régions possédant la même texture couleur dans une image donnée. Par le passé, différents algorithmes de segmentation de textures, basée sur les champs aléatoires multidimensionnels, ont été proposés pour traiter les images en niveaux de gris [3] et les images couleur [2, 4, 5].

Dans [3], les textures sont modélisées à l'aide de modèles de prédiction linéaire 2D et la segmentation exploite les erreurs de prédiction linéaire (EPL) selon l'hypothèse qu'elles sont la réalisation d'un bruit blanc gaussien. Cependant, pour les textures couleur naturelles, cette approximation sur la distribution des erreurs de prédiction linéaire a tendance à produire des résultats non-optimaux [6].

La segmentation d'images couleur texturées en utilisant les champs aléatoires de Markov est proposée dans [2]. Dans ce contexte, l'information liée à la texture est fournie par les filtres de Gabor, tandis que l'information couleur est extraite directement de l'espace CIE $L^*u^*v^*$ de chaque pixel de l'image. CIE $L^*u^*v^*$ est un modèle de représentation des couleurs développé en 1976 par la Commission internationale de l'éclairage (CIE), il caractérise une couleur à l'aide d'un paramètre d'intensité correspondant à la luminance (L) et de deux paramètres de chrominance (u et v) qui décrivent la couleur.

Le champ des classes est représenté par un mélange de densités gaussiennes multivariées. Cette hypothèse paramétrique non vérifiée produit, dans certaines situations, des résultats instables [2].

Dans cet article, nous proposons des approximations paramétriques basées sur la distribution de Wishart et les lois de mélange gaussiennes multidimensionnelles (LMGM) pour modéliser la distribution des EPL vectorielles. Ceci permettra de pallier aux points faibles d'une distribution de probabilité gaussienne certes multivariée, permettant donc de modéliser les dépendances entre les éléments des vecteurs, mais néanmoins unimodale.

La section 2 présente le modèle d'observation de l'image couleur utilisé. La section 3 est consacrée aux différents modèles utilisés pour approcher la distribution de l'EPL. L'estimation du champ des labels de classes, ou de l'image segmentée, est détaillée dans la section 4. Les sections 5 et 6 présentent respectivement les expérimentations et la conclusion.

2. Modèle d'observation de l'image.

L'étude présentée dans cet article est réalisée sur les espaces couleurs RGB (Red, Green, Blue), IHLS (Improved Hue Luminance Saturation) et L*a*b* (Luminance, Chrominance). L'information couleur, obtenue en utilisant les formules normalisées de passage RGB vers IHLS et RGB vers L*a*b*, est utilisée pour construire une image constituée de deux canaux : un canal réel pour l' 'Intensité' et un canal complexe pour la partie 'chromatique' [7, 8].

L'idée consiste à construire une image contenant deux canaux. Le premier canal correspond aux valeurs d'intensité, c'est à dire la composante 'L' des espaces IHLS et L*a*b*. Le second canal est obtenu à partir des 2 composantes chromatiques. Ces composantes sont regroupées afin d'obtenir une seule valeur complexe représentative de la chromaticité. Pour IHLS, ce deuxième canal est obtenu par l'expression suivante :

$$C = S \times \exp^{(ixH)} \quad (1)$$

où H et S sont respectivement la teinte et la saturation. Pour l'espace CIE L*a*b*, le deuxième canal est calculé par :

$$C = a^* + jb^* \quad (2)$$

où a^* et b^* sont les composantes chromatiques de l'espace CIE L*a*b*. Une image couleur peut ainsi être représentée par un vecteur défini par :

$$x_n = \begin{bmatrix} l_n \\ c_n \end{bmatrix} \quad (3)$$

où les $l_n \in \mathfrak{R}$ et les $c_n \in \mathbb{C}$, avec $n = (n_1, n_2) \in \Lambda \subset Z^2$, représentent respectivement le canal 'intensité' et le canal 'partie chromatique'. Λ est une grille finie 2D qui représente le support de l'image.

Un champ aléatoire 2D complexe multicanal peut être décrit par une famille de vecteurs aléatoires $X = (X_n)_{n \in Z^2}$ de dimension P , P représentant le nombre de canaux. La prédiction linéaire de X_n associée à un support de prédiction $D \subset \Lambda$ est définie par le vecteur :

$$\hat{X}_n = - \sum_{m \in D} A_m (X_{n-m} - \mu) \quad (4)$$

avec $\mu = E(X_n)$, $\forall n \in Z$, le vecteur moyen. Son erreur de prédiction E_n est donnée par:

$$E_n = (X_n - \mu) - \hat{X}_n. \quad (5)$$

$E = \{E_n\}_{n \in \mathbb{Z}^2}$ représente la famille de vecteurs aléatoires associés aux erreurs de prédiction. Dans le cas où le support de prédiction est causal, comme dans le cas du modèle AR QP (AutoRégressif Quart de Plan), ou AR DPNS (Autorégressif Demi-Plan Non Symétrique), la séquence d'erreur de prédiction est un bruit blanc. Tandis que pour un support de prédiction non-causal, par exemple le modèle de Gauss-Markov ou Gauss Markov Random Field (GMRF), la séquence d'erreur de prédiction est un bruit corrélé [7]. Dans la littérature, ces erreurs de prédiction linéaire ont été exploitées pour développer des méthodes de segmentation d'images texturées sur des images en niveaux de gris [3]. Pour les images couleur définies par l'équation (3), $P = 2$ pour les espaces L^*a*b^* et IHLS, tandis que pour l'espace couleur RGB, $P = 3$, avec des vecteurs de valeurs réelles, et le raisonnement reste identique. Dans cette étude, nous utilisons le modèle bidimensionnel autorégressif quart de plan (AR 2D QP) [7,8].

3. Modélisation paramétrique de la distribution de l'EPL

Dans ce papier, la méthode de segmentation des images texturées couleur est supervisée. Nous déterminons à partir d'une texture d'apprentissage, l'ensemble des paramètres du modèle, $\theta = \{\mu, \{A_m\}_{m \in D}, \Sigma_e\}$, puis nous utilisons ces paramètres pour calculer la séquence de l'EPL $e = \{e_n\}_{n \in \Lambda}$, de l'image associée à chaque texture envisagée. Classiquement, la distribution de chaque erreur de prédiction est approchée par une loi gaussienne multidimensionnelle (cf. équation 4 et 5) :

$$p(e_n | \theta) = \frac{(2\pi)^{-\omega/2}}{\sqrt{\det(\Sigma_e)}} \exp \left[-\frac{e_n^T (\Sigma_e)^{-1} e_n}{2} \right] \quad (6)$$

où $e_n = (x_n - \mu) - \hat{x}_n, n \in \Lambda$. Pour l'espace RGB, $\omega = 3$, puisque chaque vecteur d'erreur de prédiction contient 3 valeurs réelles. Dans le cas des images bi-canal complexe (équation 1 et 2) on obtient des vecteurs d'erreurs de prédiction contenant deux valeurs complexes sachant que la partie imaginaire de celle associée au canal 'Intensité' est pratiquement nulle. Ainsi, nous avons aussi, pour ces espaces, des vecteurs de taille $\omega = 3$ contenant la partie réelle du canal 'Intensité' et les parties réelles et imaginaires du canal 'partie chromatique'. Bien que cette approximation soit simple et largement utilisée, elle a deux inconvénients :

- les résultats de la segmentation ne sont ni robustes, ni stables.
- les résultats ne sont pas optimaux.

Pour éviter ces problèmes, deux autres modèles paramétriques, utilisés pour l'approximation de la distribution des EPL multi-canaux sont présentés.

3.1. Distribution de Wishart

La distribution de Wishart est une généralisation multidimensionnelle de la loi du χ^2 . Pour obtenir robustesse et stabilité, ce modèle prend en compte des observations multiples pour définir la probabilité d'une observation donnée au lieu d'utiliser une observation unique (e_n dans notre cas). Par conséquent, les vecteurs de l'EPL d'un voisinage fini sont utilisés avec le vecteur de l'EPL d'un pixel donné pour déterminer la probabilité de l'EPL de ce pixel.

Soit S_n , une matrice de α vecteurs de l'EPL de dimension ω , avec $\alpha \geq \omega$. Nous proposons de remplacer $p(e_n|\theta)$ par $p(S_n|\theta)$ dont la densité de probabilité est celle de la distribution de Wishart :

$$p(S_n|\theta) = \frac{|M_n|^{\frac{\alpha-\omega-1}{2}} \exp(-\frac{1}{2} Tr(\Sigma_e^{-1} M_n))}{2^{\alpha(\omega/2)} \pi^{\omega(\omega-1)/4} |\Sigma_e|^{\alpha/2} \prod_{j=1}^{\omega} \Gamma\left(\frac{1}{2}(\alpha+1+j)\right)} \quad (7)$$

où $M_n = S_n S_n^T$ est une matrice semi définie positive de taille $\omega \times \omega$ et Γ , la fonction gamma. Ici, nous considérons un voisinage spatial d'ordre 1 car on suppose que ce sont les voisins les plus proches de n dans Λ qui ont la probabilité la plus forte d'appartenir à la même texture. Par conséquent, α est égal à 5 et S_n est donnée par (Cf. Figure 1):

$$S_n = [e_{n-1_v}, e_{n-1_h}, e_n, e_{n+1_v}, e_{n+1_h}] \quad (8)$$

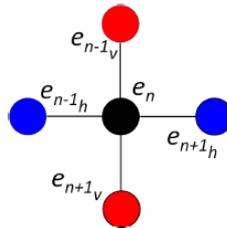


Figure 1. Voisinage Spatial d'ordre 1

3.2. Loi de mélange gaussienne multidimensionnelle

La distribution des EPL multi-canales associées à une image couleur peut également être approchée par une Loi LMGM dont la forme générale est :

$$p(e_n|\theta) = \sum_{k=1}^K \alpha_k p(e_n|\theta_k) \quad (9)$$

où les $\alpha_1, \dots, \alpha_K$ sont les probabilités a priori de chaque composante gaussienne du mélange, et $K \geq 1$ est le nombre de composantes de la LMGM, où $K = 1$ correspond à la définition d'une loi gaussienne multidimensionnelle. Chaque $\theta_k = \{\mu_{e,k}, \Sigma_{e,k}\}$ est l'ensemble des paramètres du modèle caractérisant la $k^{\text{ème}}$ composante du modèle de mélange. Ces paramètres sont estimés en utilisant l'algorithme EM (Expectation Maximization) comme dans [8]. Les valeurs de probabilité a priori doivent satisfaire :

$\alpha_k > 0, k = 1, \dots, K$ et $\sum_{k=1}^K \alpha_k = 1$, avec $k=5$ dans notre étude [6]. Du fait que les

distributions des EPL associées aux différentes classes peuvent être multimodales, les LMGM doivent permettre d'améliorer les résultats obtenus avec la loi de Gauss multidimensionnelle. Quant à la loi de Wishart, elle doit permettre de rendre plus robuste l'attribution d'une classe à un site en prenant en compte, dans la probabilité permettant de classer un site, non pas un vecteur d'EPL mais plusieurs vecteurs d'EPL.

4. Estimation du champ d'étiquettes

La première phase de la méthode de segmentation d'images texturées couleur attribue les étiquettes de classe, $\ell = \{\ell_n\}_{n \in \Lambda}$ sans réaliser de régularisation spatiale. L'attribution est réalisée selon le critère du maximum de vraisemblance qui maximise le produit des probabilités $p(e_n)$, en supposant l'indépendance des EPL les unes par rapport aux autres :

$$\hat{\ell}_n = \arg \max_{c=1, \dots, C} (p(e_n|\hat{\theta}_c)) \quad (10)$$

où C est le nombre total de classes dans l'image couleur texturée, sachant que l'ensemble des paramètres utilisé dans l'équation 10 sont ceux estimés en utilisant différentes approximation de la distribution des EPL (Cf. section 3). Au cours de la seconde phase, nous réalisons une estimation de type maximum a posteriori (MAP) pour déterminer le label (ou l'étiquette) de classe qui sera attribué à chaque pixel [2,3].

$p(l|x)$ peut s'écrire d'une manière générale sous la forme d'une distribution de Gibbs:

$$p(\ell|x) \propto \exp(-U_D(x, \ell) - U_i(\ell)) \quad (11)$$

dans laquelle U_D est une 'fonction d'énergie' dépendant du champ d'observation donné x et du champ d'étiquettes ℓ , alors que U_i est la fonction d'énergie qui dépend uniquement du champ d'étiquettes permettant la régularisation spatiale de celui-ci. U_D est calculée ainsi :

$$U_D(x, \ell) = \sum_n \left(-\log(p(e_n | \theta_{\ell_n})) \right) \quad (12)$$

où $p(e_n | \theta_{\ell_n})$ est la probabilité conditionnelle de l'EPL connaissant le label de classe de texture en n , c'est-à-dire ℓ_n .

Nous proposons d'utiliser une énergie interne au champ d'étiquettes composée de deux termes: $U_i(\ell) = U_{i,1}(\ell) + U_{i,2}(\ell)$; $U_{i,1}(\ell)$ correspond au terme d'énergie de Gibbs associé à un modèle de Potts [2,3] :

$$U_{i,1}(\ell) = \beta \left(\sum_{\langle n_1, n_2 \rangle_1} 1_{(\ell_{n_1} \neq \ell_{n_2})} + \sum_{\langle n_1, n_2 \rangle_2} 1_{(\ell_{n_1} \neq \ell_{n_2})} \right) \quad (13)$$

avec β le terme de pondération ou hyperparamètre du modèle de Potts et où $\langle n_1, n_2 \rangle_p$, $p = 1, 2$ signifie $\|n_1 - n_2\|_2 = \sqrt{p}$, $(n_1, n_2) \in \Lambda^2$, $n_1 \neq n_2$. Pour calculer $U_{i,2}(\ell)$ nous utilisons un terme d'énergie qui dépend de la taille de la région contenant uniquement des pixels connexes auxquels la même étiquette de classe a été attribuée. La taille de la région $|R_i|$, $i = 1, \dots, n_R$, où n_R est le nombre total de régions dans le champ d'étiquettes, suit une distribution de probabilité qui encourage la formation de grandes régions [9]. Son terme d'énergie est défini ainsi :

$$U_{i,2}(\ell) = \gamma \left(\sum_{i=1}^{n_R} |R_i|^{-c} \right) \quad (14)$$

où $c > 0$ est une constante qui permet de pénaliser plus ou moins les régions de petites tailles, et $\gamma \geq 0$ est un hyperparamètre qui permet de régler l'influence de cette énergie par rapport aux autres énergies (Cf. équation 11). Nous donnerons leurs valeurs dans la partie suivante, celles-ci ayant été choisies suivant [9].



Figure 2 – Base de données composée de dix images texturées

5. Expérimentations et résultats

La vérité terrain associée à la segmentation d'images naturelles complexes, en dehors d'un contexte précis, est difficile à obtenir. Sa conception est fortement influencée par la subjectivité de l'opérateur humain. Ainsi, l'évaluation de l'algorithme de segmentation d'images texturées proposé a été réalisée aussi bien sur des images naturelles de textures que sur des images de synthèse dont la vérité terrain est parfaitement connue. Les images de test proviennent de la base de textures couleur utilisées dans [10]. La base de données a été construite en utilisant des textures issues des bases Vistex et Photoshop.

Au cours de la première phase de segmentation, une seule imagerie de taille 32×32 a été utilisée pour l'apprentissage des paramètres de modèle de chaque classe. Les paramètres du modèle d'observation de l'image, l'EPL multi-canales et les jeux de paramètres pour les modèles paramétriques ont été calculés pour cette imagerie. Puis, ces paramètres sont utilisés pour calculer les classes initiales des champs d'étiquettes (Cf. équation 10) pour chacune des dix images couleur texturées représentées sur la figure 2. Des exemples de segmentations sans régularisation spatiale pour les images 3 et 10 sont représentés sur les figures 3 et 4. Dans la seconde phase de l'algorithme, la classification des régions est affinée par une régularisation spatiale en utilisant le modèle de Potts et l'énergie calculée à partir de la taille des régions (Cf. équation 11). Une méthode itérative (ICM-Iterated Conditional Mode) [11] a été utilisée pour obtenir la convergence des champs d'étiquettes. Durant cette méthode, nous utilisons β , l'hyperparamètre du modèle de Potts, comme un paramètre variant progressivement. Lors de la régularisation, le résultat de segmentation obtenu pour une valeur de β est utilisé comme champ initial pour la valeur suivante de β . Pour le terme d'énergie sur les régions, nous fixons

l'hyperparamètre γ à 2 et le coefficient c à 0.9 [9]. Les figures 3 et 4 montrent les résultats avec régularisation spatiale pour deux images texturées couleur.

Les pourcentages moyens d'erreur de classification des pixels sans régularisation spatiale ($\gamma = \beta = 0$) et avec régularisation spatiale sont présentés dans la table 1 et des courbes en fonction de β sont données sur la figure 5. L'apport des deux modèles paramétriques sur l'approximation gaussienne classique est évident. Ce résultat renforce l'hypothèse que la distribution des EPL multi-canales ne peut pas être correctement approchée avec une simple distribution gaussienne multidimensionnelle.

	Sans Régularisation			Avec Régularisation		
	RGB	IHLS	L*a*b*	RGB	IHLS	L*a*b*
Loi de Gauss	12.98	18.41	14.97	1.62	1.85	1.68
LMGM	12.47	16.83	13.67	1.41	1.58	1.52
Wishart	6.04	8.14	6.35	3.15	3.37	3.09

Tableau 1 – Pourcentage moyen d'erreur de classification des pixels des 10 textures

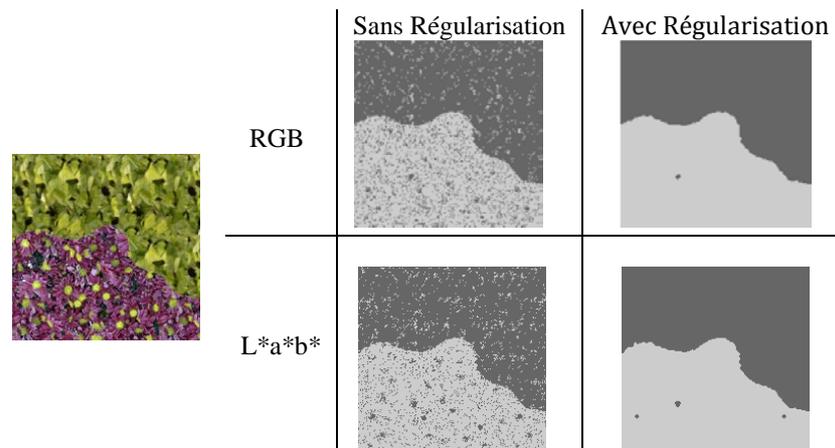


Figure 3 – Résultats de segmentation sans régularisation spatiale (colonne 'Sans Régularisation') et avec régularisation spatiale (colonne 'Avec Régularisation') pour l'image texturée couleur 3 en utilisant le modèle AR 2D QP et l'approximation LMGM pour les espaces couleurs RGB et L*a*b*.

L'approximation par LMGM de la distribution de l'EPL multi-canal donne de meilleurs résultats que les 2 autres approches en terme de pourcentage d'erreur moyen calculé après régularisation spatiale, et ce pour les dix images texturées couleur utilisées dans les expérimentations. Cependant, en considérant le même critère sans

régularisation spatiale, on observe une meilleure performance avec la distribution de Wishart. Ce résultat peut s'expliquer par le fait que la distribution de Wishart utilise plusieurs vecteurs d'EPL et donc permet une classification des sites plus robuste sans régularisation spatiale.

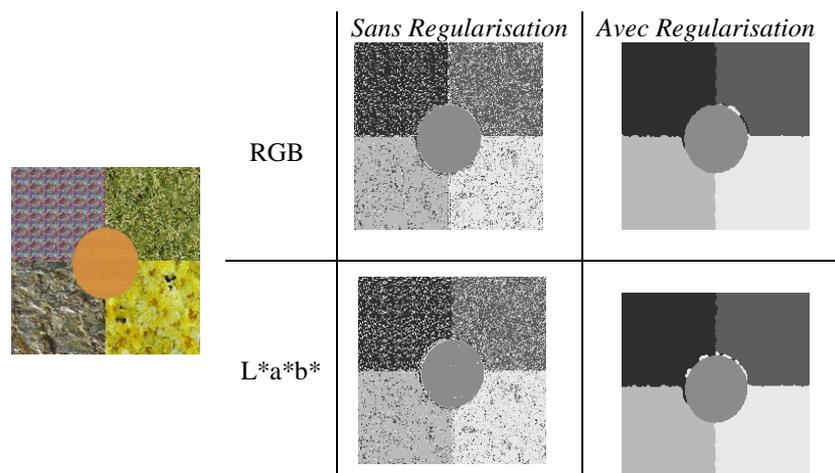


Figure 4 – Résultats de segmentation sans régularisation spatiale (colonne 'Sans Régularisation') et avec régularisation spatiale (colonne 'Avec régularisation') pour l'image texturée couleur 10 en utilisant le modèle AR 2D QP et l'approximation LMGM pour les espaces couleurs RGB et $L^*a^*b^*$.

Pour comparer les espaces couleur testés, nous traçons l'écart-type du pourcentage d'erreur (Cf. Figure 6) pour avoir un aperçu de la stabilité des résultats en fonction de l'espace couleur. Ces courbes montrent clairement des résultats meilleurs et plus stables avec les espaces couleur RGB et $L^*a^*b^*$. Cependant, l'espace couleur $L^*a^*b^*$ se montre supérieur à l'espace RGB avec une approximation gaussienne classique ou une approximation par lois de Wishart. Pour l'approximation par la LMGM de la distribution des EPL, les résultats obtenus pour les trois espaces couleur sont équivalents au niveau de la stabilité. Il est très difficile d'identifier le meilleur espace couleur.

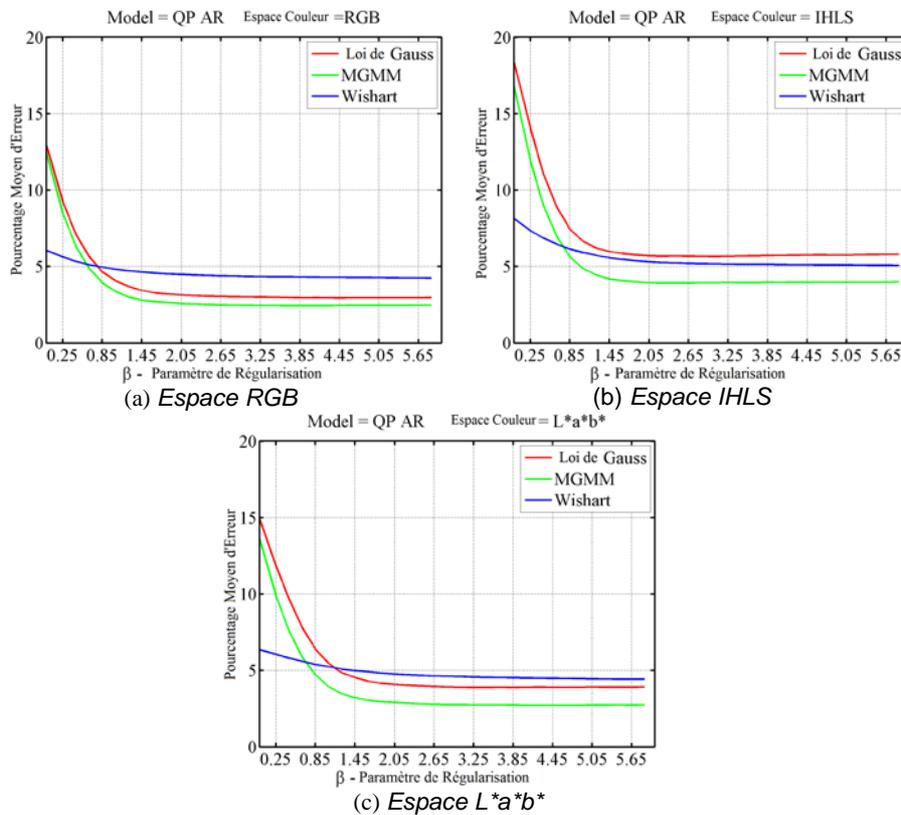


Figure 5 – Comparaison des résultats de segmentation, réalisée à l'aide du pourcentage moyen d'erreur de classification des sites, obtenus avec trois modèles paramétriques différents de distribution de l'EPL pour les espaces couleurs (a) RGB, (b) IHLS et (c) $L^*a^*b^*$.

6. Conclusion

Une comparaison de trois approches paramétriques montre qu'il est préférable d'utiliser une approximation de la distribution de l'EPL par LMGM plutôt qu'une distribution gaussienne multivariée lorsqu'un algorithme de segmentation supervisé avec régularisation spatiale est utilisé. Dans ce papier, la régularisation spatiale est réalisée à l'aide d'un modèle de Potts et d'une distribution de Gibbs attribuant une probabilité plus faible aux régions de petites tailles. L'approximation par LMGM permet ainsi d'obtenir de meilleurs résultats en segmentation d'images texturées couleur en terme de pourcentage moyen d'erreur sur la base d'images texturées couleur utilisée. Il faut aussi

noter que des résultats similaires ont été obtenus lorsque les couleurs de l'image texturée sont décrites dans l'espace RGB ou l'espace $L^*a^*b^*$. Des résultats inférieurs ont été obtenus avec l'usage de l'espace IHLS.

La distribution de Wishart produit également une bonne classification sans régularisation spatiale. Cependant, le champ initial obtenu suivant ce modèle étant déjà plus régulier qu'avec les autres modèles, la méthode itérative d'obtention du champ d'étiquettes final (algorithme ICM) ne parvient pas à produire une segmentation finale qui améliore celle obtenue avec les LMGM, l'algorithme utilisé étant dépendant de l'initialisation. Il faudrait tester alors un algorithme moins dépendant de l'initialisation comme le recuit simulé.

L'introduction d'une pénalité sur la taille des régions a permis d'améliorer nos résultats en termes de pourcentage d'erreur par rapport à ceux présentés dans [6]. De plus, cette pénalité permet une convergence plus rapide que dans le cas d'une régularisation spatiale classique utilisant uniquement le modèle de Potts.

7. Bibliographie

- [1] B. WUEST, Y. ZHANG, « Supervised Region-Based Segmentation of Quickbird Multispectral Imagery », IEEE IGARSS, vol. 2, pp 1021–1024, 2008.
- [2] Z. KATO, T.-C. PONG, « A Markov random field image segmentation model for color textured images », Image and Vision Computing, vol. 24, no 10, pp 1103–1114, 2006.
- [3] O. ALATA, C. RAMANANJARASOA, « Unsupervised textured image segmentation using 2-D quarter plane autoregressive model with four prediction supports », Pattern Recognition Letters, vol. 26, no 8, pp 1069–1081, 2005.
- [4] D. K. PANJWANI, G. HEALEY, « Markov Random Field Models for Unsupervised Segmentation of Textured Color Images », IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no 10, pp 939–954, 1995.
- [5] H. PERMUTER, J. FRANCOIS, I. JERMYN, « A study of Gaussian mixture models of color and texture features for image classification and segmentation », Pattern Recognition, vol. 39, no 4, pp 695–706, 2006.
- [6] I.-U.-H. QAZI, F. GHAZI, O. ALATA, J.-C. BURIE, C. FERNANDEZ-MALOIGNE, « A Multivariate Gaussian Mixture Model of Linear Prediction Error for Colour Texture Segmentation », EUSIPCO, pp 1537-1541, 2009.
- [7] I.-U.-H. QAZI, O. ALATA, J.-C. BURIE, C. FERNANDEZ-MALOIGNE, « Color spectral analysis for spatial structure characterization of textures in IHLS color space », Pattern Recognition, vol. 43, no 3, pp 663 - 675, 2010.
- [8] I.-U. QAZI, O. ALATA, J.-C. BURIE, A. MOUSSA AND C. FERNANDEZ-MALOIGNE 'Choice of a Pertinent Color Space for Color Texture Characterization Using Parametric Spectral Analysis' Pattern Recognition, vol. 44, pp. 16-31, 2011
- [9] Z. TU, S.-C. ZHU, « Image Segmentation by Data-Driven Markov Chain Monte Carlo », IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, pp 657–673, 2002.

[10] D.E. ILEA, P.F. WHELAN, « CTex An adaptive Unsupervised Segmentation Algorithm Based on Color-Texture Coherence », IEEE Transactions on Image Processing, vol. 17, no 10, pp 1926–1939, 2008.

[11] J. BESAG, “On the statistical analysis of dirty pictures”, Journal of the Royal Statistical Society, series B 48,259-302 (1986).

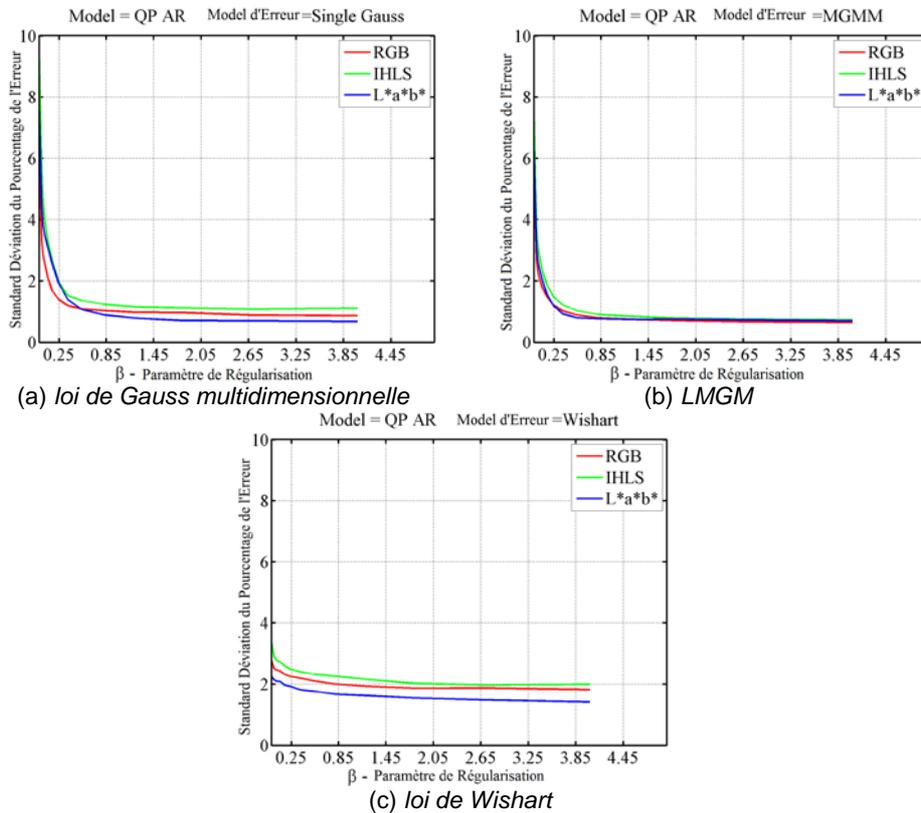


Figure 6 – Comparaison des résultats de segmentation, réalisée à l'aide de l'écart-type du pourcentage moyen d'erreur de classification des sites, obtenus pour les trois espaces couleur avec différentes approximations de la distribution de l'EPL : (a) loi de Gauss multidimensionnelle ; (b) LMGM ; (c) loi de Wishart.