

Segmentation non supervisée d'images non stationnaires avec champs de Markov évidentiels

D. BENBOUDJEMA et W. PIECZYNSKI

INT/GET, Dept. CITI, CNRS UMR 5157, 9 rue Charles Fourier, 91000 Evry Cedex

{dalila.benboudjema, wojciech.pieczynski}@int-evry.fr

Résumé – Fréquemment utilisés en traitement statistique d'images, les champs de Markov cachés (CMC) sont des outils puissants qui peuvent fournir des résultats remarquables. Cette qualité est principalement due à l'aptitude du modèle de prendre en compte des dépendances spatiales des variables aléatoires, même lorsqu'elles sont en très grand nombre, pouvant dépasser le million. Dans un tel modèle le champ caché X est supposé markovien et doit être estimé à partir du champ observé Y . Un tel traitement est possible du fait de la markovianité de X conditionnellement à Y . Ce modèle a été ensuite généralisé aux champs de Markov couples (CMC couple), où l'on suppose directement la markovianité du couple (X, Y) , qui offrent les mêmes possibilités de traitements que les CMC et permettent de mieux modéliser le bruit ce qui permet, en particulier, de mieux prendre en compte l'existence des textures. Par la suite, les CMC couples ont été généralisés aux champs de Markov triplet (CMT), où la loi du couple (X, Y) est une loi marginale d'un champ de Markov triplet $T = (X, U, Y)$, avec un champ auxiliaire U . Par ailleurs, la théorie de l'évidence peut permettre une amélioration des résultats obtenus par des traitements bayésiens dans certaines situations. Le but de cet article est d'aborder le problème de la segmentation non supervisée d'images non stationnaires en utilisant les champs de Markov évidentiels (CME), en exploitant, en particulier, un lien existant entre les CME et les CMT.

Abstract – The hidden Markov fields (HMF) are widely used in statistical image processing. Their success is mainly due to the fact that they are able to take into account the spatial dependencies, even in a very large pixel set case. In such model the hidden field X which is assumed Markovian, must be estimated from the observed field Y . Such processing is possible because the distribution X conditional on the observed process Y remains markovian. HMF has been extended to the Pairwise Markov field (PMF) model, in which one directly assumes the markovianity of the couple (X, Y) , and which offer similar processing capabilities and superior modelling ones. Afterwards, triplet Markov fields (TMF), which extend the PMF, have been proposed. In such model the distribution of the couple (X, Y) is the marginal distribution of a Markov field $T = (X, U, Y)$, where U is some auxiliary process. Otherwise, the theory of evidence can improve the bayesian processing results in some situations. The aim of this paper is to tackle the problem of the unsupervised non stationary images using the evidential Markov fields (EMF) and show the connexion between such model and the TMF one.

1. Introduction

Les champs de Markov sont des outils probabilistes assez largement utilisés en traitement d'images [7, 11, 20]. Leur intérêt majeur est de pouvoir modéliser les dépendances spatiales des variables aléatoires, dont les réalisations modélisent les quantités observées ou recherchées, de façon relativement simple et se prêtant bien aux divers traitements. Divers modèles fondés sur les champs markoviens ont été proposés, parmi lesquels les champs de Markov cachés (CMC), très utilisés en segmentation d'images. Ceux-ci consistent à considérer deux champs stochastiques $X = (X_s)_{s \in S}$ et $Y = (Y_s)_{s \in S}$, avec S ensemble des pixels, dans lesquels chaque variable aléatoire X_s est à valeurs dans l'ensemble fini des classes $\Omega = \{\omega_1, \dots, \omega_k\}$ et chaque Y_s est à valeurs dans l'ensemble des nombres réels R . En segmentation d'images, le problème est d'estimer le champ inobservable directement (ou « caché ») $X = x$ à partir du champ observé $Y = y$. Dans le modèle CMC classique X est de Markov et la forme du bruit (loi de Y conditionnelle à X) est simple, ce qui implique la markovianité de X *a posteriori* (conditionnelle à Y). Cette dernière markovianité permet alors de mettre en œuvre les différents traitements bayésiens [7, 11, 20]. Cependant, cette même simplicité est difficile à

justifier et peut poser problème dans certaines situations, comme la présence des textures [10]. Dans le modèle « champ de Markov couple » (CMC couple), plus général que CMC dans la mesure où X n'est plus nécessairement de Markov, on suppose directement la markovianité de $Z = (X, Y)$ [14, 15]. Le champ Y est alors markovien conditionnellement à X (ce qui autorise une modélisation plus complexe des bruits), et le champ X est markovien conditionnellement à Y (ce qui autorise les mêmes traitements que dans les CMC). Par la suite, les CMC couple ont été généralisés aux « champs de Markov triplet » (CMT) dans lesquels on introduit un troisième champ aléatoire $U = (U_s)_{s \in S}$, où chaque v.a. U_s est à valeurs dans l'ensemble $\Lambda = \{\lambda_1, \dots, \lambda_M\}$, et l'on suppose la markovianité du triplet $T = (X, U, Y)$ [13]. Le modèle CMT est alors plus général que les CMC couple car la loi de $Z = (X, Y)$, qui est une loi marginale de celle de $T = (X, U, Y)$, n'est plus nécessairement markovienne. Cependant, lorsque le cardinal de l'ensemble Λ n'est pas trop élevé, les CMT autorisent encore les traitements bayésiens permettant d'estimer $X = x$ à partir de $Y = y$ [4, 5, 13]. Notons que Λ peut avoir une signification concrète ou pas. A titre d'exemple, Λ peut modéliser l'existence des

sous-classes (voir [5]), ou encore l'existence des diverses stationnarités du champs des classes X (voir [4]).

Par ailleurs, la théorie de l'évidence [1, 17, 18] peut permettre une amélioration des résultats obtenus par des traitements bayésiens dans un certain nombre de situations [1]. Cette théorie est particulièrement bien adaptée à la fusion des informations fournies par des capteurs de nature différentes comme, à titre d'exemple, un capteur optique et un capteur radar dans une image satellitaire (voir [19]). Cependant, elle a été très peu utilisée, jusqu'à présent, dans le contexte des champs de Markov cachés [2, 6, 16].

L'objet de cet article est de montrer l'intérêt de l'introduction des fonctions de masse au niveau de la loi *a priori* d'un champs de Markov caché en segmentation non supervisée d'images non stationnaires. En procédant de manière proche à celle utilisée dans le contexte des chaînes de Markov cachées (voir [8, 9]), nous utilisons le fait que l'introduction d'un certain type de l'affaiblissement au niveau de la loi *a priori* aboutit, formellement, à un champ de Markov triplet. Le problème de l'estimation des paramètres dans les champs triplet étant proche de celui se posant dans les champs couple, nous adaptions la méthode proposée dans [3, 5]. Les résultats de diverses expérimentations, dont une est présentée en section 5, attestent de l'intérêt de la méthode proposée.

2. Théorie de l' évidence

La théorie de l' évidence a été introduite en 1967 par Dempster puis reprise par Shafer en 1976 [17]. Parfois décrite comme une généralisation de la théorie probabiliste classique, elle a l' avantage de pouvoir représenter de manière parfois plus complète l' incertitude sur un événement. Elle peut également permettre une amélioration des résultats obtenus par des traitements bayésiens dans un certain nombre de situations. Cela est en particulier vrai lorsque les paramètres du modèles sont mal connus, auquel cas on remplace certaines lois de probabilité, par exemple la loi *a priori* de X , par des « fonctions de masse » évidentielles. Une telle fonction de masse M est une application de $P(\Omega)$ (ensemble des parties de Ω) dans $[0,1]$ vérifiant $M(\emptyset)=0$ et $\sum_{A \in P(\Omega)} M(A)=1$. Notons qu'une fonction de masse M définit

les « plausibilités » $Pl(A) = \sum_{A \cap B \neq \emptyset} M(B)$ et les « crédibilités »

$Cr(A) = \sum_{B \subseteq A} M(B)$. Une des possibilité de définition d'une

fonction de masse à partir d'une loi de probabilité est l'utilisation d'un « affaiblissement ». Cette démarche peut être utilisée pour la loi *a priori* de X , lorsque cette dernière est connue trop imparfaitement. La loi *a posteriori* de X est alors obtenue par la fusion de Dempster-Shafer de la masse évidentielle obtenue par affaiblissement avec une loi de probabilité définie par les observations $Y = y$. Cette démarche n'est pas immédiate dans le cas des modèles de Markov cachés car la fusion de Dempster-Shafer détruit la markovianité ; cependant, il a été montré que le résultat de cette fusion est la loi marginale d'un modèle « triplet », ce qui autorise les traitements d'intérêt [12]. En particulier, lorsque la loi *a priori* est non stationnaire et inconnue, son estimation donnera nécessairement une loi erronée et son affaiblissement

peut éventuellement présenter un intérêt. En effet, cet intérêt a été montré dans le cas des chaînes de Markov cachées dans [8, 9], et nous montrons ci-dessous qu'il peut également exister dans le cas des champs de Markov.

3. Champs de Markov évidentiels

Soit S un ensemble de pixels, X un champ de Markov où chaque v.a X_s est à valeur dans un ensemble fini de classes $\Omega = \{\omega_1, \dots, \omega_k\}$, avec la loi $p(x) \propto \exp[-W(x)]$ donnée par l'énergie :

$$W(x) = \sum_{(s,t) \in C_H} \alpha_H (1 - 2\delta(x_s, x_t)) + \sum_{(s,t) \in C_V} \alpha_V (1 - 2\delta(x_s, x_t)) \quad (1)$$

où C_H (respectivement C_V) est l'ensemble des cliques d'ordre deux horizontales (respectivement verticales), et $\delta(x_s, x_t)$ vérifie $\delta(x_s, x_t) = 1$ pour $x_s = x_t$, et $\delta(x_s, x_t) = 0$ pour $x_s \neq x_t$. Avec les hypothèses classiques portant sur le bruit, la loi du couple (X, Y) est donnée par :

$$p(x, y) = \gamma \exp[-W(x) + \sum_{s \in S} \text{Log}(p(y_s | x_s))] \quad (2)$$

Le modèle ainsi défini est un champ de Markov caché classique.

Plaçons nous maintenant dans le cadre évidentiel. Dans les champs de Markov évidentiels (CME), on introduit un champ $U = (U_s)_{s \in S}$ où chaque U_s est à valeurs dans l' ensemble des parties de Ω notée $\Lambda = P(\Omega)$. La fonction de masse M_1 est un champ de Markov évidentiel (CME) défini sur $(P(\Omega))^n$ (avec $n = \text{Card}(S)$) si :

$$M_1(A) = \gamma \exp\left(-\sum_{c \in C} \psi_c(A_c)\right) \quad (3)$$

où $A = (A_s)_{s \in S}$ et $A_c = (A_s)_{s \in c}$, l' ensemble des cliques C correspondant à un voisinage donné.

En posant $M_2 \propto \prod_{s \in S} p(y_s | x_s)$ la probabilité sur Ω^n définie à partir du processus observé $Y = y \in R^n$ ($y = (y_s)_{s \in S}$ est fixé), la probabilité $M = M_1 \oplus M_2$ est la probabilité conditionnelle $p(x|y)$, avec $p(x, y)$ la loi marginale d'un CMT dont l'énergie est donnée comme suit :

$$\begin{aligned} \varphi_c(t_c) &= \varphi_c(x_c, u_c, y_c) \\ &= \begin{cases} \psi_c(u_c) & \text{pour } \text{Card}(c) > 1 \\ \psi_c(u_c) + \log(p(y_s | x_s)) & \text{pour } c = \{s\} \end{cases} \end{aligned} \quad (4)$$

En effet, $M = M_1 \oplus M_2$ s' écrit

$$\begin{aligned} M(x) &= (M_1 \oplus M_2)(x) \\ &\propto \sum_{x \in \Omega^n} \left[\exp\left(-\sum_{c \in C} \psi_c(u_c)\right) \prod_{s \in S} p(y_s | x_s) \right] \\ &\propto \sum_{x \in \Omega^n} \left[\exp\left(-\sum_{c \in C} \psi_c(u_c) + \sum_{s \in S} \log(p(y_s | x_s))\right) \right] \end{aligned} \quad (5)$$

Notons que dans la somme figurant dans (5) $x = (x_s)_{s \in S}$ est fixé et $u = (u_s)_{s \in S}$ varie dans l' ensemble $[P(\Omega)]^n$. Plus

précisément, l'énergie dans (5) est définie pour $(x, u) \in \Delta^n$ avec $\Delta \subset \Omega \times \Lambda$ défini par $\Delta = \{(\omega, A) | \omega \in A\}$.

A titre d'exemple, pour un champ de Markov à deux classes on a: $\Omega = \{\omega_1, \omega_2\}$, $\Lambda = P(\Omega) = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}\}$, et le couple (X_s, U_s) prend ses valeurs dans $\{(\omega_1, \{\omega_1\}), (\omega_1, \{\omega_1, \omega_2\}), (\omega_2, \{\omega_2\}), (\omega_2, \{\omega_1, \omega_2\})\}$. La fonction de masse $M(x)$ devient:

$$M(x) = \sum_{u | (x, u, y) \in \Delta^n \times R^n} \left[\exp - \left(\sum_{c \in C} \varphi_c(u_c) - \sum_{s \in S} \log(p(y_s | x_s)) \right) \right] \quad (6)$$

$$= \sum_{u | (x, u, y) \in \Delta^n \times R^n} \exp - \varphi(x_c, u_c, y_c)$$

Ainsi on retrouve l'expression du champ Triplet $T = (X, U, Y)$ donné comme suit :

$$p(x, u, y) = \gamma \exp[-W(x, u) + \sum_{s \in S} \text{Log}(p(y_s | x_s))] \quad (7)$$

On peut ainsi dire que le résultat de la fusion de Dempster - Shafer est la loi conditionnelle $p(x|y)$ induite par un champ Triplet $T = (X, U, Y)$, avec le processus auxiliaire U à valeurs dans l'ensemble des parties $P(\Omega)$ de Ω . D' autre part, pour $x = u$ l'affaiblissement est réduit à néant, on retrouve l'énergie $W(x)$ classique du champ de Markov caché (équation (1)).

Remarque

Notons que la théorie de l'évidence peut également être utilisée au niveau des capteurs, comme proposé dans [2], l'introduction des champs de Markov triplet. Supposons que l'on a toujours deux classes et l'on souhaite utiliser le modèle classique donné par (1) et (2) en présence des nuages. Le champs de Markov $p(x) \propto \exp[-W(x)]$ n'est pas affecté par les nuages ; par contre, la probabilité $M_2 \propto \prod_{s \in S} p(y_s | x_s)$ change de nature et tout se passe comme s'il y avait trois classes : ω_1 , ω_2 , et « nuages ». La troisième classe « nuages » peut alors être interprétée comme une classe n'apportant aucune information d'intérêt et assimilée à Ω . Finalement, la probabilité $M_2 \propto \prod_{s \in S} p(y_s | x_s)$ fournie par les capteurs peut être interprétée comme une masse définie sur $\{\{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}\}^n$ et l'on montre que sa fusion avec le champ de Markov $p(x) \propto \exp[-W(x)]$ reste un champ de Markov. Ce dernier apparaît alors immédiatement comme une extension du champs de Markov *a posteriori* classique et peut être classiquement utilisé à des fins de segmentation.

4. Apprentissage

Dans cette section nous décrivons brièvement une méthode, proche de celle proposée dans [3], d'estimation des paramètres du modèle. Elle utilise moindres carrés pour l'estimation des paramètres du champ de Markov (U, X) , et des méthodes classiques utilisant les moments pour l'estimation des paramètres du bruit. Comme dans [3] et afin de rendre cette méthode automatique l'estimation conditionnelle itérative (ECI) est utilisée. L'ECI est une procédure générale d'estimation des paramètres dans le cas des données cachées [15]. Soit X et Y deux champs aléatoires dont la loi conjointe dépend d'un paramètres θ , et

soit $\hat{\theta} = \hat{\theta}(X, Y)$ un estimateur de celui ci à partir des données complètes. Le champ X étant inobservable, l'idée de l'ECI est d'approcher $\hat{\theta} = \hat{\theta}(X, Y)$ par son espérance conditionnelle à $Y = y$. Or, l'espérance conditionnelle de $\hat{\theta}$ notée $E[\hat{\theta}|Y]$, dépend du paramètre θ . Ainsi, pour calculer $E[\hat{\theta}|Y]$, on est amené à considérer la valeur θ_q (valeur courante) du paramètre. La valeur suivante est donnée par $\theta_{q+1} = E_{\theta_q}[\hat{\theta}|Y]$. La procédure itérative ECI est donc :

- On se donne une valeur θ_0
- On calcule la valeur θ_{q+1} (valeur suivante) à partir de θ_q (valeur courante) et de $Y = y$ par :
$$\theta_{q+1} = E_q[\hat{\theta}(X, Y) | Y = y] \quad (8)$$

où E_q est l'espérance conditionnelle correspondant à θ_q .

$E_q[\hat{\theta}|Y = y]$ n'est pas explicitement calculable, mais la simulation des réalisations de X à partir de la loi *a posteriori* est possible. Alors on utilise l'approximation suivante :

$$\theta_{q+1} = \frac{1}{r} E_q \left[\hat{\theta}(x_1, y) + \dots + \hat{\theta}(x_r, y) \right] \quad (9)$$

Où x_1, x_2, \dots, x_r sont les r simulations des réalisations de X selon la loi conditionnelle à Y .

5. Expérimentation

Considérons une image non stationnaire à deux classes illustrée en FIG. 1 (a) et sa version bruitée (b), avec un bruit blanc gaussien de variance 1 et de moyennes 0 et 2. L'utilisation de la méthode bayésienne Maximum des Marginales a Posteriori (MPM) fondée sur le modèle affaiblit triplet (5) donne 9.37% de taux d'erreur (image (d)), alors que celle fondée sur le modèle CMC classique (1), (2) donne 22.1% de taux d'erreur (image (c)). Dans les deux cas, la segmentation est non supervisée dans laquelle les paramètres ont été initialisés en utilisant la méthode des k - means et estimés par la méthode de type ECI décrite ci-dessus. Les paramètres ainsi obtenus sont représentés dans TAB. 1.

TAB. 1 : valeurs des paramètres estimés

Paramètres	CMC	CME
μ_1	0.89	-0.2
μ_2	2.05	1.99
σ_1	1.38	0.87
σ_2	0.99	0.99
α_{1H}, α_{1V}	0.64, 0.64	0.24, 0.28
α_{2H}, α_{2V}		0.35, 0.26
α_{3H}, α_{3V}		0.00, 0.02

TAB. 1 montre que les paramètres du bruit estimés à partir du modèle évidentiel sont plus proches des paramètres réels que ceux estimés à partir de CMC classiques. Par ailleurs, aucune conclusion ne peut être tirée quant à l'estimation des paramètres d'interaction α_i étant donnée que leurs vraies valeurs sont inconnues. Nous pouvons cependant observer que l'utilisation des CME a permis l'amélioration du résultat

de segmentation non supervisée, et donc conjecturer que ce dernier modèle représente mieux l'image considérée.

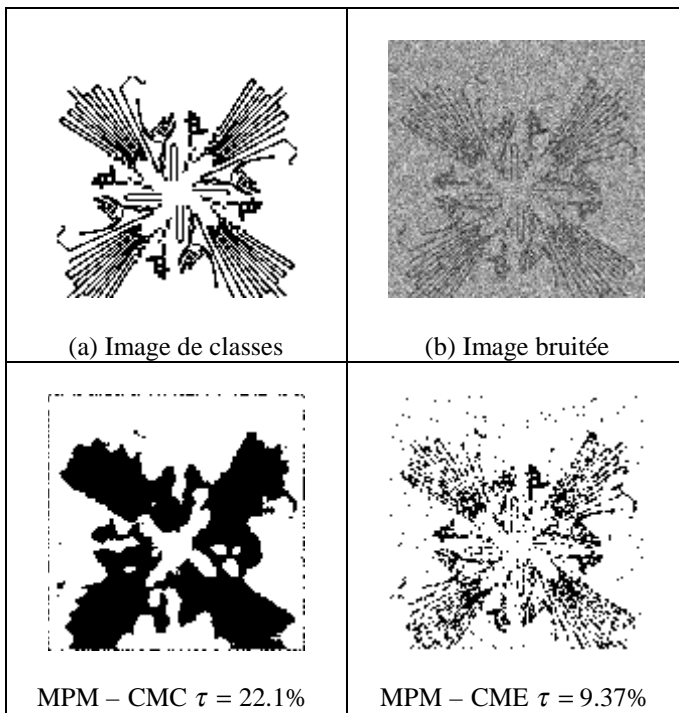


FIG. 1 : segmentation non supervisée

6. Conclusion

Dans cet article nous avons étudié la possibilité d' utiliser la théorie de l' évidence dans le cadre des champs de Markov cachés (CMC). A cet effet, nous avons proposé un modèle d'affaiblissement, aboutissant à un champ de Markov triplet, ainsi qu'une méthode d'estimation de paramètres proche de celle décrite dans [3]. L' intérêt de l' affaiblissement de la loi *priori* d'un champs de Markov caché en segmentation non supervisée d'images non stationnaires a également été validé par le biais d'un exemple.

Comme perspectives, nous pouvons envisager l' extension de ce modèle aux champs de Markov couple, ou l' utilisation d' un bruitage plus complexe. L'extension au cas multivarié peut également être envisagée.

Références

- [1] A. Appriou, Probabilités et incertitude en fusion de données multisenseurs, *Revue Scientifique et Technique de la Défense*, No. 11, pp. 27-40, 1991.
- [2] A. Bendjebbour, Y. Delignon, L. Fouque, V. Samson, and W. Pieczynski, Multisensor images segmentation using Dempster-Shafer fusion in Markov fields context, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 39, No. 8, pp. 1789-1798, 2001.
- [3] D. Benboudjema and W. Pieczynski, Parameter estimation in pairwise Markov fields, *Advanced Concepts for Intelligent Vision Systems (ACIVS 04)*, Aug. 31-Sept. 3, Brussels, Belgium, 2004.
- [4] D. Benboudjema and W. Pieczynski, Segmenting non stationary images with triplet Markov fields,

- International Conference on Image Processing (ICIP 2005), Genova, Italy, September 11-14, 2005.
- [5] D. Benboudjema and W. Pieczynski, Unsupervised image segmentation using triplet Markov fields, *Computer Vision and Image Understanding*, to appear, 2005.
- [6] S. Foucher, M. Germain, J.-M. Boucher, and G. B. Benié, Multisource classification using ICM and Dempster-Shafer theory, *IEEE Trans. on Instrumentation and Measurement*, Vol. 51, No. 2, pp. 277-281, 2002.
- [7] S. Geman and D. Geman, *Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*, *IEEE Trans. On PAMI*, vol. 6, pp. 721-741, 1984.
- [8] P. Lanchantin et W. Pieczynski, Chaînes et arbres de Markov évidentiels avec applications à la segmentation des processus non stationnaires, *Traitement du Signal*, Vol. 22, No. 1, pp. 15-26, 2005.
- [9] P. Lanchantin and W. Pieczynski, Unsupervised restoration of hidden non stationary Markov chain using evidential priors, *IEEE Trans. on Signal Processing*, to appear, 2005.
- [10] D. Melas and S. P. Wilson, Double Markov Random Fields and Bayesian Image Segmentation, *IEEE Trans. on Signal Processing*, Vol. 50, No. 2, pp. 357-365, 2002.
- [11] P. Pérez. Markov random fields and images. *CWI Quarterly*, Vol. 11, No. 4, pp. 413-437, 1998.
- [12] W. Pieczynski and D. Benboudjema, Triplet Markov fields and theory of evidence, submitted to *Image and Vision Computing Journal*, October 2003.
- [13] W. Pieczynski, D. Benboudjema, and P. Lanchantin, Statistical image segmentation using Triplet Markov Fields, *SPIE's International Symposium on Remote Sensing*, September 22-27, Crete, Greece, 2002.
- [14] W. Pieczynski and A.-N. Tebbache, Pairwise Markov random fields and segmentation of textured images, *Machine Graphics and Vision*, Vol. 9, No. 3, pp. 705-718, 2000.
- [15] W. Pieczynski, Modèles de Markov en traitement d' images *Traitement du Signal*, Vol. 20, No. 3, pp. 255-278, 2003.
- [16] A. Sankar, A. Banerjee, N. Banerjee, S. Brahma, B. Kartikeyan, M. Chakraborty, and K. L. Majumder, Landcover classification in MRF context using Dempster-Shafer fusion for multisensor imagery, *IEEE Trans. on Image Processing*, Vol. 14, No. 5, pp. 634-645, 2005.
- [17] G. Shafer, A mathematical theory of evidence, *Princeton University Press*, Princeton, 1976.
- [18] P. Smets, Belief functions versus probability functions, In B. Bouchon, L. Saitta, and R. R. Yager (eds), Uncertainty and intelligent systems, *Berlin, Springer-Verlag*, pp. 17-24, 1988.
- [19] F. Tupin, H. Maitre, and I. Bloch, A first step toward automatic interpretation of SAR images using evidential fusion of several structure detectors, *IEEE Trans. on GRS*, Vol. 37, No. 3, pp. 1327-1343, 1999.
- [20] G. Winkler, *Image analysis, random fields and and Markov Chain Monte Carlo Methods: a mathematical introduction*, Springer, 2003.