

Approximation gaussienne de champs de Markov pour la segmentation non supervisée d'images

Sonia OUALI¹, Jean-Baptiste COURBOT¹, Romain PIERRON², Olivier HAEBERLE¹

¹IRIMAS UR 7499, ²LVBE UR 3391
Université de Haute-Alsace, Mulhouse, France
sonia.ouali@uha.fr

Résumé – Nous nous intéressons au problème de segmentation bayésienne d'images. Les méthodes avec a priori markoviens étant très coûteuses calculatoirement, nous proposons une approximation via un modèle de champ gaussien. Nous introduisons le modèle et la démarche d'inférence associée pour la segmentation non supervisée, et les résultats numériques montrent l'intérêt en performance, et en temps de calcul, de notre approche.

Abstract – We address the problem of Bayesian image segmentation. Methods based on Markov field prior being computationally demanding, we propose an approximation through a Gaussian random field model. We present the model as well as the corresponding inference framework, and numerical results show the advantages both in performances and in computation times of the proposed method.

1 Introduction

Nous nous intéressons ici au problème de segmentation bayésienne d'image. Malgré l'efficacité des modèles de champ de Markov [1, 7] pour la segmentation, leur utilisation est coûteuse calculatoirement. Pour y remédier, nous proposons d'utiliser un champ gaussien [10] pour approcher un champ de Markov.

1.1 Champs de Markov et segmentation d'image

Soit \mathbf{X} un champ de Markov de réalisation $\mathbf{x} \in \Omega^{|\mathcal{S}|}$, avec Ω l'ensemble des classes envisagées, et \mathcal{S} la grille sur laquelle les sites (pixels) de l'image sont localisés. On choisit pour \mathbf{x} un modèle de Potts :

$$p(\mathbf{x}) \propto \exp(-\alpha \sum_{(s,t) \in E} (1 - 2\delta(x_s, x_t))) \quad (1)$$

avec α la granularité de \mathbf{x} , E l'ensemble des paires de pixels voisins dans l'image et δ la fonction de Kronecker.

Soit $\mathbf{Y} = \mathbf{y}$ l'image observée. Supposons un bruit gaussien indépendant, dépendant des classes de \mathbf{x} :

$$p(\mathbf{y}|\mathbf{x}) = \prod_{s \in \mathcal{S}} p(y_s|x_s) \text{ et } p(y_s|x_s = \omega_k) \sim \mathcal{N}(y_s; \mu_k, \sigma_k) \quad (2)$$

Ainsi avec (1)(2), la loi jointe du modèle de champs de Markov cachés (CMC) est :

$$p(\mathbf{x}, \mathbf{y}) \propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x}) = p(\mathbf{x}) \prod_{s \in \mathcal{S}} p(y_s|x_s). \quad (3)$$

Deux estimateurs sont courants dans la segmentation par CMC. Le *Maximum A Posteriori (MAP)* peut être approché à

l'aide de l'algorithme *Iterative Conditional Mode (ICM)* [1] :

$$\hat{\mathbf{x}}_s^{\text{MAP}} = \arg \max_{x \in \Omega^N} p(\mathbf{X} = \mathbf{x} | \mathbf{Y} = \mathbf{y}). \quad (4)$$

Le *Marginal Posterior Mode (MPM)* peut quant à lui être approché par l'algorithme de Marroquin [9] :

$$\forall s \in \mathcal{S} : \hat{x}_s^{\text{MPM}} = \arg \max_{\omega \in \Omega} p(x_s = \omega | \mathbf{Y} = \mathbf{y}). \quad (5)$$

Le calcul direct de $p(\mathbf{x}|\mathbf{y})$ est impossible, car calculer $p(\mathbf{x})$ nécessite le calcul d'une fonction de partition sur $\Omega^{|\mathcal{S}|}$. Il est donc nécessaire de recourir à des méthodes approchées, comme l'échantillonneur de Gibbs qui permet de simuler \mathbf{x} en suivant une approche de Monte Carlo par chaîne de Markov (MCMC), ce qui est très lent. Une solution partielle à cette lenteur est l'utilisation d'un échantillonneur chromatique [8] qui parallélise ce qui peut l'être dans l'échantillonneur de Gibbs classique.

1.2 Champs de Gauss-Markov

Soit $\mathbf{Z} \in \mathbb{R}^S$ un champ gaussien (GMRF pour *Gaussian Markov Random Field*) et \mathbf{z} sa réalisation :

$$\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (6)$$

avec $\mathbf{x}, \boldsymbol{\mu} \in \mathbb{R}^S$, $\boldsymbol{\Sigma} \in \mathbb{R}^{S \times S}$ la matrice de covariance associée ; et la matrice de précision $\mathbf{Q} = \boldsymbol{\Sigma}^{-1}$.

Dans le cas générique, simuler et inférer \mathbf{z} est limité par la taille de $\boldsymbol{\Sigma}$. C'est pourquoi il est nécessaire de faire l'hypothèse que $\boldsymbol{\Sigma}$ est circulante. Cela implique que \mathbf{z} est un GMRF stationnaire, mais aussi que l'image est sur un tore.

Ainsi, supposons qu'une observation \mathbf{y} est formée par une version bruitée d'un champ gaussien :

$$\mathbf{y} = \mathbf{z} + \boldsymbol{\epsilon}, \quad (7)$$

avec \mathbf{z} un GMRF et $\epsilon \sim \mathcal{N}(\mathbf{0}, \Sigma_\epsilon)$. Alors la distribution *a posteriori* est aussi une loi normale :

$$p(\mathbf{Z} = \mathbf{z} | \mathbf{Y} = \mathbf{y}) = \mathcal{N}(\mathbf{z}; \boldsymbol{\mu}^*, \mathbf{Q}^*), \text{ où} \quad (8)$$

$$\begin{cases} \boldsymbol{\mu}^* &= \mathbf{Q}^{*-1}(\Sigma_\epsilon^{-1} \boldsymbol{\mu} + \Sigma_\epsilon^{-1} \mathbf{Y}) \\ \mathbf{Q}^* &= \Sigma_\epsilon^{-1} + \Sigma_\epsilon^{-1} \end{cases}$$

Ainsi $\boldsymbol{\mu}^*$ est aussi l'estimateur du MAP de \mathbf{z} .

Nous supposons de plus que Σ est paramétrées par une fonction de corrélation ρ , c'est-à-dire que pour toute paire de pixels \mathbf{a}, \mathbf{b} dans l'image :

$$\Sigma_{\mathbf{a}, \mathbf{b}} = \sigma_z^2 \rho(\mathbf{a}, \mathbf{b}; r) = \sigma_z^2 \exp\left(-\frac{1}{r_z^d} \|\mathbf{a} - \mathbf{b}\|_t^d\right) \quad (9)$$

où $\|\mathbf{a} - \mathbf{b}\|_t$ représente la distance entre \mathbf{a} et \mathbf{b} sur le tore. Dans la suite, on prendra soit $d = 1$ (corrélation exponentielle) ou $d = 2$ (gaussienne). Notons que cette fonction de corrélation pourrait prendre d'autres formes (Matérn par exemple).

\mathbf{Q} et Σ sont circulantes, elles sont donc caractérisées par leur base, respectivement notées $\mathbf{b}_\mathbf{Q} \in \mathbb{R}^S$ et $\mathbf{b}_\Sigma \in \mathbb{R}^S$. Cela permet d'effectuer les calculs suivants :

- $\mathbf{b}_\mathbf{Q} = \text{IDFT}(1 \div \text{DFT}(\mathbf{b}_\Sigma))$, avec DFT et IDFT la transformée de Fourier discrète et son inverse, et \div la division élément par élément.
- soit $\mathbf{z} \in \mathbb{R}^S$ un vecteur : $\mathbf{z}\mathbf{Q} = \text{IDFT}(\text{DFT}(\mathbf{z}) \otimes \text{DFT}(\mathbf{b}_\mathbf{Q})) = \mathbf{z} * \mathbf{b}_\mathbf{Q}$, avec \otimes la multiplication élément par élément, et $*$ l'opérateur de convolution.

Ainsi, ni \mathbf{Q} , ni Σ ne doivent être stockées car $\mathbf{b}_\mathbf{Q}$ et \mathbf{b}_Σ permettent de faire tous les calculs dans le domaine de Fourier.

Enfin, notons que les GMRF sont également markoviens. En effet, la fonction de corrélation ρ (9) implique qu'au-delà d'une certaine distance entre deux sites \mathbf{a} et \mathbf{b} , la corrélation entre $z_\mathbf{a}$ et $z_\mathbf{b}$ devient négligeable. Cela permet d'associer ρ à un voisinage, et d'énoncer une propriété de markovianité. Ce choix de ρ permet considérer que les GMRF sont aussi markoviens.

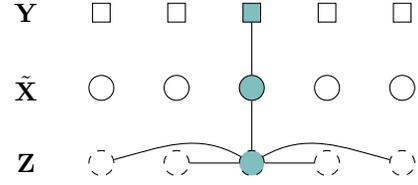
Dans cet article, nous montrons comment la similarité des deux modèles peut être exploitée pour proposer une segmentation d'image correspondant à un champ de Markov en utilisant les facilités de calculs offertes par les champs gaussiens. La Section 2 détaille le modèle proposé et la méthode d'inférence associée, et les résultats numériques sont présentés en Section 3.

2 Modèle, approximation, algorithmes

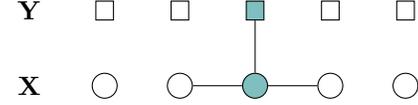
2.1 Modèle

Le modèle exact de la loi jointe avec a priori de Markov est résumé dans l'équation (3). Ici, nous proposons un modèle approché, qui ne représente pas explicitement la markovianité du processus recherché. On notera $\tilde{\mathbf{X}}$ le processus caché du modèle approché pour le distinguer de \mathbf{X} du modèle CMC.

Nous utilisons pour notre modèle deux champs latents : $\tilde{\mathbf{X}} \in \Omega^{|\mathcal{S}|}$ avec $\Omega \stackrel{\text{def.}}{=} \{\omega_0, \omega_1\}$, forme la segmentation binaire que l'on souhaite déterminer à partir de l'observation $\mathbf{Y} = \mathbf{y}$. Ce champ est lié à un champ gaussien $\mathbf{Z} = \mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}_z, \Sigma)$, de



(a) Graphe du modèle proposé. Les liens entre les éléments de \mathbf{Z} sont régis par Σ et donc par ρ , qui décrit la portée des corrélations (Eq. 9).



(b) Graphe pour le modèle CMC classique.

FIGURE 1 – Graphes de dépendances des distributions étudiées, conditionnellement à un site s (en bleu). Une absence de lien représente une indépendance conditionnelle des deux variables.

sorte à avoir $\forall s \in \mathcal{S}$:

$$\tilde{x}_s = \psi(z_s) \stackrel{\text{def.}}{=} \begin{cases} \omega_1 & \text{si } z_s > \mu_z. \\ \omega_0 & \text{sinon.} \end{cases} \quad (10)$$

$\mathbf{Y} = \mathbf{y}$ est décrit par rapport à $\tilde{\mathbf{X}} = \tilde{\mathbf{x}}$ comme dans (2). La Fig. 1 représente le graphe de dépendance du modèle proposé.

Ainsi, la loi jointe du modèle est :

$$\begin{aligned} p(\mathbf{y}, \tilde{\mathbf{x}}, \mathbf{z}) &= p(\mathbf{y} | \tilde{\mathbf{x}}) p(\tilde{\mathbf{x}} | \mathbf{z}) p(\mathbf{z}) \\ &= p(\mathbf{z}) \prod_{s \in \mathcal{S}} p(y_s | \tilde{x}_s) p(\tilde{x}_s | z_s) \end{aligned} \quad (11)$$

2.2 Approximation

Comme illustré en Fig. 1 et dans les équations (3)(11), les distributions jointes des modèles exacts et approchés diffèrent ; ce sera donc également le cas pour les lois *a posteriori*.

Une possibilité serait de formaliser cette différence, et de la minimiser, avec des outils d'inférence variationnelle [2]. Mais ceux-ci nécessitent le calcul de la distribution d'origine $p(\mathbf{x}, \mathbf{y})$, qui est précisément inaccessible en raison de la fonction de partition.

Pour pouvoir approcher $p(\mathbf{x} | \mathbf{y})$ par $p(\tilde{\mathbf{x}} | \mathbf{y})$, nous faisons l'hypothèse suivante :

$$\mathbf{y} \simeq \mathbf{z} + \mathbf{b}, \text{ et } \mathbf{b} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}\sigma_b^2), \quad (12)$$

avec $\sigma_b > 0$. Cela permet de se placer, pour \mathbf{z} , dans le cadre des équations (7)(8). Cela requiert du modèle exact qu'il y ait :

1. bruit indépendant, sans quoi, il faut considérer une matrice de covariance au lieu de σ_b .
2. un bruit gaussien pour la formation des images (2).

Nous pouvons aussi formuler un critère du MAP pour $\tilde{\mathbf{x}}$:

$$\tilde{\mathbf{x}}^{\text{MAP}} = \arg \max_{\tilde{\mathbf{x}} \in \Omega^S} p(\tilde{\mathbf{x}} | \mathbf{z}) p(\mathbf{z} | \mathbf{y}) = \psi \left(\arg \max_{\mathbf{z} \in \mathbb{R}^S} p(\mathbf{z} | \mathbf{y}) \right), \quad (13)$$

avec $\psi(\mathbf{z}) = \{\psi(z_1), \dots, \psi(z_{|\mathcal{S}|})\}$, et $p(\mathbf{z} | \mathbf{y})$ donné par (8).

Finalement, les paramètres du modèle sont (2)(9)(10)(12) :

$$\Theta = \{\mu_0, \mu_1, \sigma_0, \sigma_1, \mu_z, \sigma_z, r_z, \sigma_b\}. \quad (14)$$

2.3 Estimation des paramètres avec EM et SEM

La démarche proposée suppose qu'il existe un Θ permettant d'avoir une approximation $\tilde{\mathbf{x}}^{\text{MAP}}$ la meilleure possible du réel \mathbf{x} . Pour cela, nous proposons d'utiliser une approche basée sur les algorithmes EM ou son variant stochastique, SEM.

Plusieurs articles de la littérature comparent les deux approches [5, 6]. Il est généralement admis que, en dépit de son aspect stochastique, SEM est plus rapide et peut sortir des maxima locaux qui peuvent être atteints dans EM, pour lequel la convergence vers un maximum local est en revanche établie. Nous étudions ici les deux possibilités.

Stochastic Expectation-Maximization (SEM). L'algorithme SEM [3] consiste à alterner simulations des variables latentes et ré-estimations des paramètres du modèle. Dans notre modèle, il se traduit par la répétition à l'étape q de :

- Simulation de $\mathbf{z}^{(q)} \sim p_{\Theta^{(q-1)}}(\mathbf{z}|\mathbf{y})$, et $\tilde{\mathbf{x}}^{(q)} = \psi(\mathbf{z}^{(q)})$
- Estimation de $\Theta^{(q)}$ à l'aide de $\mathbf{z}^{(q)}$, $\tilde{\mathbf{x}}^{(q)}$, \mathbf{y}

Ce dernier point est géré comme suit :

- les estimateurs de μ_0 , μ_1 , σ_0 , σ_1 sont obtenus par les estimateurs du maximum de vraisemblance (EMV) sur $(\tilde{\mathbf{x}}, \mathbf{y})$ comme dans un modèle classique. De même σ_z et σ_b sont obtenus par EMV avec (\mathbf{z}, \mathbf{y})
- Il n'existe pas d'EMV pour r (9), cette valeur est donc obtenue par moindres carrés.
- Enfin, pour « localiser » \mathbf{z} en intensité à mi-chemin entre les intensités des deux classes représentées dans $\tilde{\mathbf{x}}$, nous prenons $\mu_z = (\mu_0 + \mu_1)/2$.

Algorithme Expectation-Maximization. L'algorithme EM [4] repose sur le calcul des probabilités *a posteriori* en chaque site de l'image, et en la pondération des EMV par ces probabilités. Ces probabilités n'étant pas accessibles directement, nous recourons à une approche par échantillonnage : en réalisant suffisamment de simulations de $\mathbf{z} \sim p(\mathbf{z}|\mathbf{y})$, les probabilités *a posteriori* sont approchées par l'estimateur des fréquences. Ainsi, notre approche de l'algorithme EM est, à l'itération q :

- Simuler N réalisations de $\mathbf{z} \sim p_{\Theta^{(q-1)}}(\mathbf{z}|\mathbf{y})$, afin de calculer $\hat{p}(z_s = \omega_k | \mathbf{y})$, $\forall s \forall k$ avec les fréquences empiriques.
- Calculer $\mathbf{z}^{(q)} = \arg \max_{\mathbf{z} \in \mathbb{R}^S} p_{\Theta^{(q-1)}}(\mathbf{z}|\mathbf{y})$: équation (8).
- Estimer $\Theta^{(q)}$ avec les EMV pondérés par les $\hat{p}(z_s = \omega_k | \mathbf{y})$ lorsque les EMV sont disponibles, par moindres carrés pour r , et avec $\mu_z = (\mu_0 + \mu_1)/2$.

Pour EM, comme pour SEM, nous n'utilisons pas les estimateurs usuels avec lesquels ces méthodes ont été proposés. Ainsi dans la suite nous les appellerons *quasi-EM* (qEM) et *quasi-SEM* (qSEM).

Finalement, la démarche pour la segmentation d'image est :

- 1) Estimer Θ par qEM ou qSEM,
- 2) Effectuer le calcul au sens du MAP de \mathbf{z} (8),
- 3) Appliquer ψ pour obtenir une estimation de $\tilde{\mathbf{x}}$.

3 Résultats numériques

Pour valider notre approche, nous proposons trois études numériques sur données de synthèse.

Expérience 1. Nous comparons les différentes possibilités d'approximation en contexte non supervisé : algorithme qEM ($N = 100$ répétitions pour l'échantillonnage) ou qSEM, et fonction de corrélation gaussienne ($d = 2$ dans (9)) et exponentielle ($d = 1$). Ces méthodes sont comparées sur des images simulées correspondant au modèle CMC original (3), avec des variations de $\Delta\mu = \mu_1 - \mu_0$ et de $\sigma_0 = \sigma_1$. Nous mesurons également le rapport signal à bruit (RSB) comme $\text{RSB} = 20 \log_{10}(\Delta\mu/2\sigma_0)$.

Les résultats sont présentés en Fig. 2. Ils mettent en évidence que la corrélation exponentielle permet une meilleure segmentation, quel que soit le RSB. En revanche, la différence de résultat entre qEM et qSEM est assez faible, hormis à très faible RSB pour lesquels il est préférable d'utiliser qEM (2% d'erreur en moins). Comme qEM est plus lent (d'un facteur 2.8 en moyenne) que qSEM, il est en général préférable d'utiliser qSEM.

Expérience 2. Nous étudions ensuite comment se comporte la méthode proposée (dans le cas $d = 1$, $\text{RSB} = 0$ dB) en fonction de la taille de l'image considérée, notamment concernant les temps de calculs de la segmentation seule. Nous comparons la segmentation par la méthode proposée (en utilisant (8) et Θ estimé par qSEM), la segmentation par l'algorithme de Marroquin [9], par ICM [1], reposant chacun sur un parcours point à point ou chromatique [8] de l'image. Deux constats sont permis à l'examen de la Fig. 3a :

- Les méthodes approchant le MAP par ICM sont peu efficaces, quoique rapide pour la version chromatique. La version optimale pour le modèle exact est l'algorithme de Marroquin reposant sur un parcours chromatique. La méthode que nous proposons est aussi efficace en terme de taux d'erreur que ce dernier, avec un temps de calcul 53 fois moindre en moyenne.
- La méthode approchée montre, en non supervisé, des résultats aussi bons que la méthode exacte supervisée.

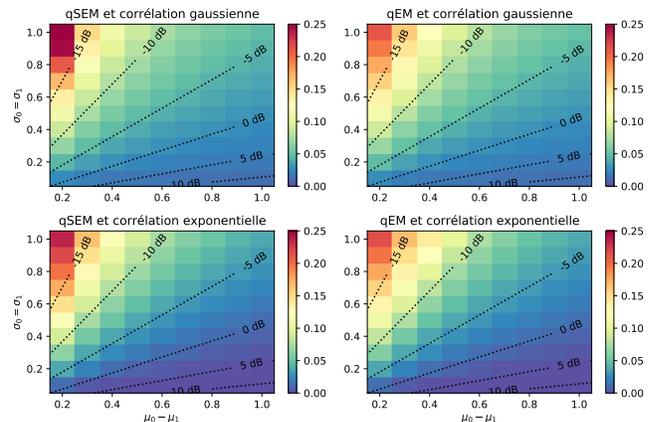
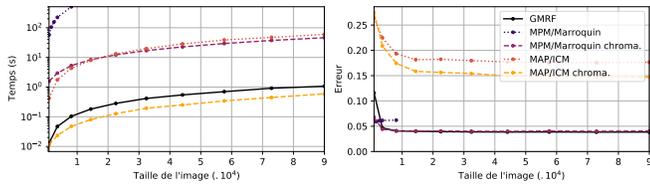
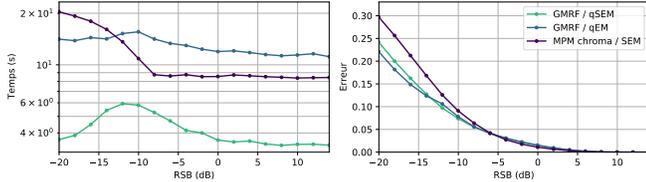


FIGURE 2 – Résultats pour l'exp. 1, dans lesquels chaque valeur est une moyenne sur 50 réalisations.



(a) Résultats pour l'exp. 2 : segmentation supervisée d'images. Le modèle non chromatique pour l'algorithme de Marroquin étant très lent, seuls quelques points sont montrés.



(b) Résultats pour l'exp. 3 : segmentation non supervisée d'images.

FIGURE 3 – Résultats numériques pour les expériences 2 et 3. À gauche : temps de calculs, et à droite : taux d'erreur. Chaque point est obtenu par moyenne sur 100 réalisations.

Expérience 3. Finalement, nous comparons les taux d'erreurs dans un contexte non supervisé pour notre méthode (avec qEM et qSEM) avec, pour le modèle exact, l'algorithme SEM pour l'estimation des paramètres et l'algorithme de Marroquin chromatique pour la segmentation. La figure 3b montre que notre méthode surpasse ce dernier à faible RSB, et lui est équivalente à RSB favorable. Par ailleurs l'utilisation de qSEM fait diminuer le temps de calcul d'un facteur 3.1 en moyenne, et qEM est encore plus robuste à bruit élevé (2% de gain). D'autres exemples (non montrés) suggèrent une robustesse de la méthode proposée à une disproportion des classes, et à une asymétrie du bruit ($\sigma_0 \neq \sigma_1$).

La figure 4 illustre des résultats sur données simulées et réelles. Ces dernières sont obtenues lors de l'examen de bois de vigne au microscope de fluorescence, et montrent les alvéoles du bois. La méthode proposée permet de mettre en évidence les alvéoles via z^{MAP} et de proposer, in fine, une segmentation très proche de celle obtenue par l'algorithme de Marroquin (3% de différence entre les deux résultats) pour un temps de calcul 2,5 fois moindre.

4 Discussion

Les expériences numériques montrent de très bons résultats à la fois en eux-mêmes, et par comparaison avec les méthodes classiques de segmentation bayésienne dans un modèle CMC. Aucune des deux approches n'est en pratique exacte : dans les méthodes usuelles, le MAP comme le MPM sont approchés, et le résultat d'un échantillonneur de Gibbs est approché car en temps fini. Ainsi, l'approche proposée consiste à déplacer l'approximation pour avoir un modèle approché, et une inférence exacte sur celui-ci. Les résultats obtenus montrent que, numériquement, les deux approches sont équivalentes à RSB

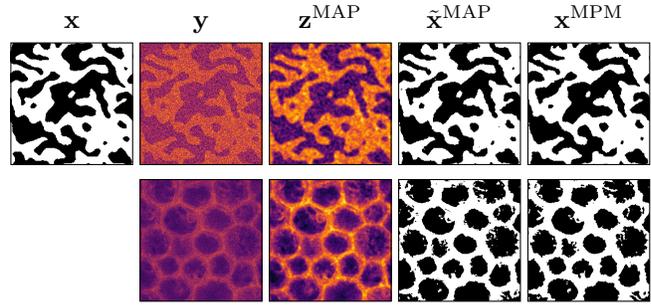


FIGURE 4 – Exemple de résultats sur image de synthèse (-6 dB, première ligne) et réelle (seconde ligne). La méthode proposée est illustrée par z^{MAP} et \tilde{x}^{MAP} (3.90% d'erreur pour l'image de synthèse), et x^{MPM} est obtenu par l'algorithme de Marroquin (3.99% d'erreur).

favorable, et que notre rapproche est plus efficace à RSB défavorable. De nombreuses investigations restent à mener, notamment pour consolider ces premiers résultats, et les étendre aux cas des images couleurs, multi-dimensionnelles, et multi-classes.

Références

- [1] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society : Series B (Methodological)*, 48(3) :259–279, 1986.
- [2] D. M. Blei, A. Kucukelbir, and J. D. McAuliffe. Variational inference : A review for statisticians. *Journal of the American statistical Association*, 112(518) :859–877, 2017.
- [3] G. Celeux, D. Chauveau, and J. Diebolt. Stochastic versions of the EM algorithm : an experimental study in the mixture case. *Journal of statistical computation and simulation*, 55(4) :287–314, 1996.
- [4] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society : Series B (Methodological)*, 39(1) :1–22, 1977.
- [5] J. G. Dias and M. Wedel. An empirical comparison of EM, SEM and MCMC performance for problematic Gaussian mixture likelihoods. *Statistics and Computing*, 14(4) :323–332, 2004.
- [6] S. Faria and G. Soromenho. Comparison of EM and SEM algorithms in Poisson regression models : a simulation study. *Communications in Statistics-Simulation and Computation*, 41(4) :497–509, 2012.
- [7] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6) :721–741, 1984.
- [8] J. Gonzalez, Y. Low, A. Gretton, and C. Guestrin. Parallel Gibbs sampling : From colored fields to thin junction trees. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 324–332. JMLR Workshop and Conference Proceedings, 2011.
- [9] J. Marroquin, Q. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. *Journal of the american statistical association*, 82(397) :76–89, 1987.
- [10] H. Rue and L. Held. *Gaussian Markov random fields : theory and applications*. Chapman and Hall/CRC, 2005.