TOMOGRAPHIC IMAGE RECONSTRUCTION WITH A SPATIALLY VARYING GAUSSIAN MIXTURE PRIOR

Katerina Papadimitriou, Christophoros Nikou

Department of Computer Science and Engineering,
University of Ioannina, 45110 Ioannina, Greece
{apapadim,cnikou}@cs.uoi.gr

ABSTRACT

A spatially varying Gaussian mixture model (SVGMM) prior is employed to ensure the preservation of region boundaries in penalized likelihood tomographic image reconstruction. Spatially varying Gaussian mixture models are characterized by the dependence of their mixing proportions on location (*contextual mixing proportions*) and they have been successfully used in image segmentation. The proposed model imposes a Student's t-distribution on the local differences of the *contextual mixing proportions* and its parameters are automatically estimated by a variational Expectation-Maximization (EM) algorithm. The tomographic reconstruction algorithm is an iterative process consisting of alternating between an optimization of the SVGMM parameters and an optimization for updating the unknown image using also the EM algorithm. Numerical experiments on various photon limited image scenarios show that the proposed model is more accurate than the widely used Gibbs prior.

Index Terms— Emission tomography, iterative image reconstruction, expectation-maximization (EM) algorithm, spatially varying Gaussian mixture models (GMM), Student's t-distribution, edge preservation.

1. INTRODUCTION

Maximum a posteriori (MAP) or penalized maximum likelihood tomographic reconstruction methods have gained increasing acceptance in the last three decades [1]. These methods impose a prior probability density function (pdf) on the image to be reconstructed which usually aims to encourage the image to be smooth so as to suppress the effect of noise. This assumption is based on the knowledge that, because of its blurring effect, the system (projection) matrix suppresses image detail. Therefore any such detail present in the reconstruction is more probably to have arisen from noise.

A common model for the prior is the Markov random field (MRF) expressed by the Gibbs distribution and many methods were proposed in that framework, differing on the choice of the potential function [2, 3, 4]. The total variation has also been used as a smoothing prior for suppressing unstable oscillations around the image edges [5, 6].

The notion of clustered intensity histogram was introduced in a penalized likelihood method in [7], where the unknown image is modeled by a mixture of Gamma distributions enforcing positivity. A monotonically decreasing surrogate objective function resulting in a closed form expression is proposed in [8] while the median root prior was also used to impose spatial smoothness and stabilize the solution [9]. Also, a nonlocal prior was designed [10] where the definition of a pixel's neighborhood is broadened. Finally, image priors based on Gaussian mixtures whose parameters are estimated using variational-Bayes methodology were proposed for image segmentation and restoration [11] and tomographic reconstruction [12].

In this paper, we propose a MAP tomographic reconstruction algorithm based on a spatially varying Gaussian mixture model (SVGMM) image prior [13], which is appropriate for both emission and transmission tomography. Contrary to the Gibbs prior, which is determined by a parameter that has to be fixed in advance, in spatially varying mixtures, the model parameters are automatically estimated from the image. This yields location-dependent smoothing which cannot be modeled by the Gibbs distribution. Therefore, the advantage of this type of model is twofold. At first, edges are better preserved than in the standard Gibbs prior due to their implicit modeling and the clustering assumption increases the interaction between similar in intensity pixels. In the related literature, many spatially varying Gaussian mixtures have been proposed and could be applied as priors in tomography [13, 14, 15, 16]. For the proof of concept, we have applied the model proposed in [13] because it models image edges as a continuous line process whose parameters provide a continuous (analog) edge map.

The image reconstruction process integrates this prior in a standard MAP-EM iterative algorithm, where the SVGMM parameters and the unknown image are estimated in an alternating scheme. Numerical experiments using photon-limited images reveal the supremacy of the method with respect to the widely used Gibbs prior.

2. THE IMAGE MODEL

Let $\mathbf{f}$ be the vectorized form of the image to be reconstructed. Let also $\mathbf{g}$ be the observed projections (sinogram), also in vectorized form and let $\mathbf{H}$ represent the projection matrix. Penalized likelihood models rely on the stochastic interpretation of Tikhonov regularization by introducing an appropriate prior $p(\mathbf{f})$ for the image $\mathbf{f}$. The likelihood function $p(\mathbf{g}|\mathbf{f})$ is related to the posterior probability $p(\mathbf{f}|\mathbf{g})$ by the Bayes rule $p(\mathbf{f}|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f})p(\mathbf{f})$.

In tomography, the likelihood $p(\mathbf{g}|\mathbf{f})$ is a Poisson distribu-

tion assuming independence between counts

$$p(\mathbf{g}|\mathbf{f}) = \prod_{n=1}^{N} ([\mathbf{H}\mathbf{f}]_n)^{g_n} \frac{\exp(-[\mathbf{H}\mathbf{f}]_n)}{g_n!}, \quad (1)$$

where N is the number of projection measures, g_n is the n^{th} component of $\mathbf{g}$ and $[\mathbf{H}\mathbf{f}]_n$ is the n^{th} component of vector $\mathbf{H}\mathbf{f}$. MAP estimates for the image $\mathbf{f}$ may be obtained by maximizing the log-posterior:

$$\log p(\mathbf{f}|\mathbf{g}) = \log p(\mathbf{g}|\mathbf{f}) + \log p(\mathbf{f}) \quad (2)$$

with respect to $\mathbf{f}$.

Based on the hypothesis that the image to be reconstructed should be smooth, the general assumption is that $p(\mathbf{f})$ represents a Markov random field represented by a Gibbs distribution and its variants. Here, we propose to model the image by a spatially varying Gaussian mixture model (SVGMM) which has shown to be very effective for image segmentation [13]. It differs from the standard GMM [17] in the definition of the mixing proportions. More precisely, in the SVGMM, each pixel $\mathbf{f}_n$, $n = 1, ..., N$ has a distinct vector of mixing proportions denoted by π_j^n, $j = 1, ..., J$, with J being the number of Gaussian kernels. We call these parameters *contextual mixing proportions* to distinguish them from the mixing proportions of a standard GMM. Hence, the probability of a distinct pixel is expressed by:

$$p(\mathbf{f}_n; \pi, \mu, \Sigma) = \sum_{j=1}^{J} \pi_j^n \mathcal{N}(\mathbf{f}_n; \mu_j, \Sigma_j) \quad (3)$$

where $0 \leq \pi_j^n \leq 1, \quad \sum_{j=1}^{J} \pi_j^n = 1$ for $j = 1, 2, ..., J$ and $n = 1, 2, ..., N$, $\mathcal{N}(\cdot)$ is the Gaussian distribution, μ_j are the Gaussian means and Σ_j are the Gaussian covariance matrices. Hence, the probability of the image is computed by assuming pixel independence, which is common in modeling images by mixtures of distributions:

$$p(\mathbf{f}) = \prod_{n=1}^{N} \sum_{j=1}^{J} \pi_j^n \mathcal{N}(\mathbf{f}_n; \mu_j, \Sigma_j) \quad (4)$$

Apart from enforcing pixel clustering, this prior preserves the edges in the image because the local differences of the *contextual mixing proportions* are considered to follow a univariate Student's t-distribution. Following the definition of the Student's t-distribution [18], a two step generative model provides the clique potential functions:

$$\pi_j^n - \pi_j^k \sim \mathcal{N}(0, \beta_{jd}^2 / u_j^{nk}), \quad (5)$$

$$u_j^{nk} \sim \mathcal{G}(\nu_{jd}/2, \nu_{jd}/2), \ \forall n, j, d, \ k \in \gamma_d(n), \quad (6)$$

where $\mathcal{G}(\cdot)$ is the Gamma distribution, $\gamma_d(n)$ is the set of neighbors of the pixel indexed by n, with respect to the d^{th} adjacency type (e.g. horizontal, vertical, diagonal). This model first draws u_j^{nk} from a Gamma distribution parameterized by ν_{jd} and then considers that the local differences of the mixing proportions follow a Gaussian distribution with zero mean and standard deviation β_{jd}^2 / u_j^{nk}.

This generative model, whose graphical representation is shown in Fig. 1, allows clustering of the image pixels around the Gaussian means and imposes edge preservation through the Student's t-distribution of the mixing proportions. More specifically, as $u_j^{nk} \to +\infty$ the distribution tightens around zero, and enforces neighboring *contextual mixing proportions* to be smooth. On the other hand, when $u_j^{nk} \to 0$ the distribution tends to be uninformative, and enforces no smoothness. Consequently, the variables u_j^{nk} provide a very detailed description of the boundary structure of the image. Estimation of model parameters through a standard MAP-EM approach is intractable due the comp[lexity of the model and the suitable framework is provided by variational inference which yields estimates for all of the parameters of the model [13].

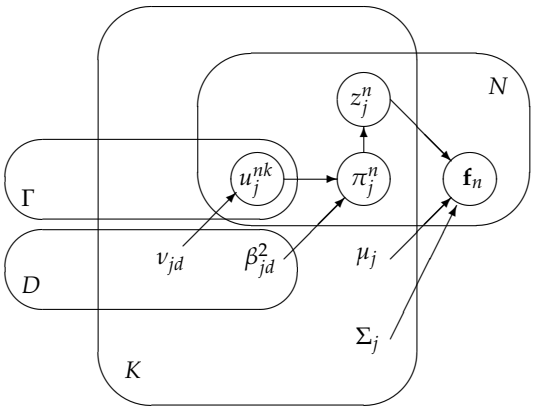


Fig. 1. Graphical model for the edge preserving model. Superscripts and subscripts $n, k \in [1, N]$ denote pixel index, subscript $j \in [1, K]$ denotes segment index, $d \in [1, D]$ describes the neighborhood direction type. Γ equals the maximum number of possible neighbors.

The maximization of (2) with respect to $\mathbf{f}$ may be accomplished by the maximum *a posteriori* expectation-maximization (MAP-EM) algorithm that provides the update for the intensity of the j^{th} pixel of the unknown image [1, 2]:

$$\hat{\mathbf{f}}_j^{(t+1)} = \frac{\hat{\mathbf{f}}_j^{(t)}}{\sum_{i'} \mathbf{H}_{i'j} + \frac{\partial \log p(\mathbf{f})}{\partial \mathbf{f}}\Big|_{\mathbf{f}=\hat{\mathbf{f}}_j^{(t)}}} \sum_i \mathbf{H}_{ij} \frac{\mathbf{g}_i}{\sum_k \mathbf{H}_{ik} \hat{\mathbf{f}}_k^{(t)}}. \quad (7)$$

This one-step-late EM (OSL-EM) algorithm evaluates the derivative term in (7) using the previous image estimate. Other methods, such as the conjugate gradient descent may be employed but they also reach a local maximum of (2) as the mixture model is not convex with respect to $\mathbf{f}$.

The overall algorithm consists of an alternating optimization scheme. One step consists in estimating the parameters of the SVGMM using the EM algorithm [13] with the image $\mathbf{f}$ being fixed. Having the SVGMM parameters fixed from the first step, the second step consists in estimating $\mathbf{f}$ by the OSL-EM algorithm update (7). Algorithm 1 summarizes the different steps. The algorithm stops when the estimated image does not change significantly or when a predefined number of iterations is reached, which is common in this type of alternating optimization methods [7, 9, 10].

Algorithm 1 MAP-EM tomographic reconstruction using a SVGMM prior

input: A sinogram $\mathbf{g}$, a threshold ϵ, MAXiterations.
output: The unknown image $\mathbf{f}$.
Initialize $\mathbf{f}$ by an image with constant intensity.
counter=0.
while $||\mathbf{f}^{(t+1)} - \mathbf{f}^{(t)}|| > \epsilon$ **and** counter $\leq$ MAXiterations **do**
  Estimate the parameters $\{\mu_j, \Sigma_j, \pi_j^n, \beta_{jd}, \nu_{jd}, u_j^{nk}\}$, for $n, k \in [1, N]$, $j \in [1, K]$ and $d \in [1, D]$ of the SVGMM in (4) using the EM algorithm as described in [13].
  Estimate the image $\mathbf{f}$ using the update in (7).
  counter++
end while

3. EXPERIMENTAL RESULTS

The performance of the proposed tomographic reconstruction model was examined using the well known Shepp-Logan phantom. We have set $J = 5$ clusters for the mixture model taking into account the segments of the phantom. The algorithm stopped when $\epsilon = 10^{-3}$ or when 60 iterations were reached. The method was evaluated with respect to the standard maximum likelihood EM (ML-EM), the established MAP-EM algorithm with a Gibbs prior [2] and a version of the proposed method using a standard GMM as prior. A number of performance indices were used. To this end, degraded images were generated from the initial image by modifying the total photon counts. More specifically, as the initial phantom has approximatively 75 counts per pixel, images having 75, 55, 35 and 15 photons/pixel on average were generated to degrade the signal quality.

At first, the algorithms were put in test in terms of the improvement in signal to noise ratio (ISNR) with respect to a reconstruction obtained by a simple filtered back-projection using the Ram-Lak filter:

$$\text{ISNR} = 10 \log_{10} \frac{||\mathbf{f} - \mathbf{f}_{FBP}||^2}{||\mathbf{f} - \hat{\mathbf{f}}||^2} \tag{8}$$

where $\mathbf{f}$ is the ground truth image, $\mathbf{f}_{FBP}$ is the reconstructed image by filtered back-projection and $\hat{\mathbf{f}}$ is the reconstructed image using the proposed image model. Practically, ISNR measures the improvement (or deterioration) in the quality of the reconstruction of the proposed method with respect to the reconstruction obtained by filtered back-projection. Moreover, the consistency of the method was measured by the bias (BIAS) and the variance (VAR) of the reconstructed images:

$$\text{BIAS} = ||\mathbf{f} - \bar{\mathbf{f}}||, \quad \text{VAR} = \sum_{k=1}^{M} ||\bar{\mathbf{f}} - \hat{\mathbf{f}}_k||^2, \tag{9}$$

with

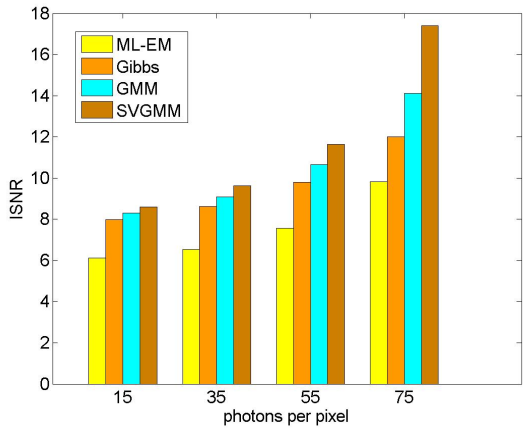
$$\bar{\mathbf{f}} = \frac{1}{M} \sum_{k=1}^{M} \hat{\mathbf{f}}_k, \tag{10}$$

where $\mathbf{f}$ is the ground truth image and $\hat{\mathbf{f}}_k$, for $k = 1, ..., M$, is the k^{th} reconstructed image, obtained from $M = 40$ different realizations for each noise level. Finally, we also included in the evaluation the structural similarity index (SSIM) [19], which represents the visual distortion between the ground truth and the reconstructed image:

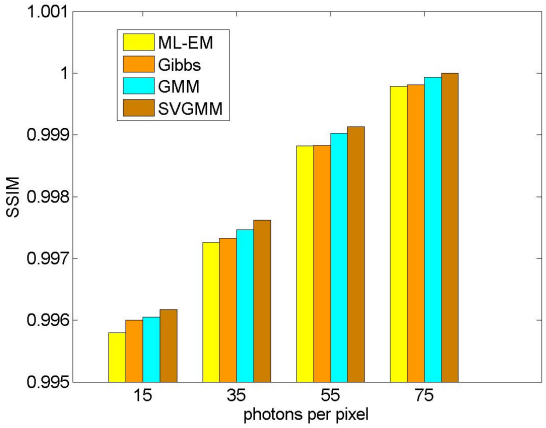
$$\text{SSIM}(\mathbf{f}, \hat{\mathbf{f}}) = \frac{(2\mu_{\mathbf{f}}\mu_{\hat{\mathbf{f}}} + C_1)(2\sigma_{\mathbf{f}\hat{\mathbf{f}}} + C_2)}{(\mu_{\mathbf{f}}^2 + \mu_{\hat{\mathbf{f}}}^2 + C_1)(\sigma_{\mathbf{f}}^2 + \sigma_{\hat{\mathbf{f}}}^2 + C_2)}, \tag{11}$$

where $\mu_{\mathbf{f}}$ and $\mu_{\hat{\mathbf{f}}}$ denote the mean intensity of the ground truth and the estimated image, $\sigma_{\mathbf{f}}$ and $\sigma_{\hat{\mathbf{f}}}$ are the standard deviations of the two images, $\sigma_{\mathbf{f}\hat{\mathbf{f}}}$ is the covariance of $\mathbf{f}$ and $\hat{\mathbf{f}}$ and C_1 and C_2 are constants added to avoid instability. The above statistics are calculated locally on equally sized windows centered at each image pixel and the average values over all pixels are reported here.

The statistical comparisons for these indices are shown in Fig. 2 and Fig. 3. For the ISNR and the SSIM (Fig. 2), their mean values over the $M = 40$ experiments are shown. All of the obtained ISNR and SSIM values are very close to these mean values as their standard deviations over the whole set of experiments are very small. For the bias and variance, the values obtained by (9) are shown. As it can be observed



(a)



(b)

Fig. 2. (a) ISNR and (b) structural similarity of the compared methods (mean values over 40 experiments).

in these figures, as the noise decreases (the number of photon counts per pixel increases) the ISNR becomes larger with

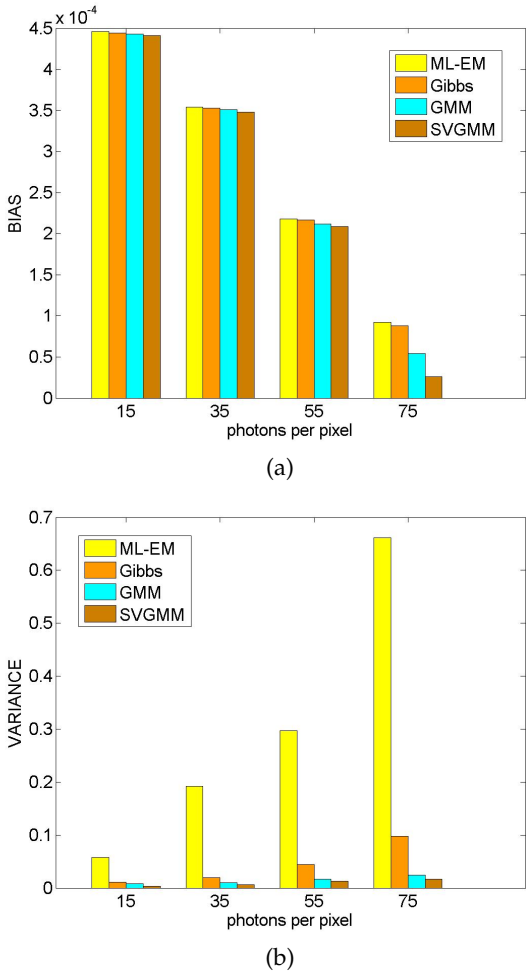


Fig. 3. (a) Bias and (b) variance of the compared methods.

the SVGMM method providing better results. Similar numerical results are obtained for the peak SNR (PSNR). The same stands for the bias which decreases when the noise decreases as we have more pure data and the obtained mean estimates are closer to the ground truth as the photon counts increase. Furthermore, the variance of the estimates is relatively consistent for the SVGMM which is due to the clustering effect of the prior (notice this effect also for the simple GMM). In all these indices, the SVGMM prior clearly shows a better performance with respect to the other priors.

In general, the image provided by the SVGMM prior is sharper. This is also confirmed by a visual inspection through scaling (zooming) of the estimated images in Fig. 4. Finally, to highlight the accuracy of the proposed model, the estimated image intensities along a scan line are shown in Fig. 5 for the Gibbs and the SVGMM priors, where it can be seen the SVGMM model provides values which are closer to the ground truth. The execution of the algorithm takes on average 4 minutes on a standard PC using MATLAB without any optimization.

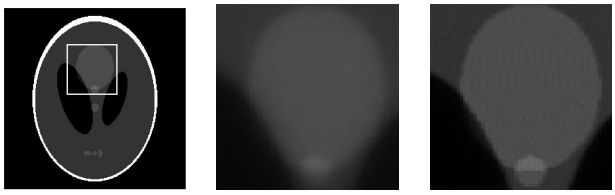


Fig. 4. (a) Original image and a zoomed region inside it. The zoomed region reconstructed using (b) the Gibbs prior and (c) the SVGMM prior with 75 counts per pixel.

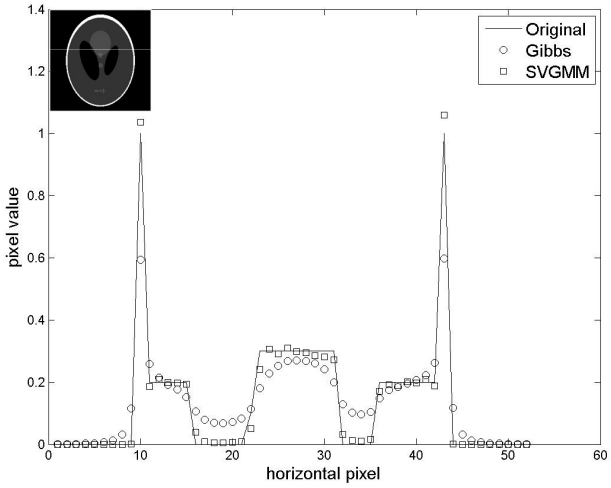


Fig. 5. Comparison of horizontal profiles between the original and the reconstructed images provided by the proposed SVGMM and the Gibbs prior 75 counts per pixel.

4. CONCLUSION

In this paper, an iterative reconstruction algorithm based on a spatially varying Gaussian mixture model was proposed. The main contribution of this work is the effectiveness and robustness of the prior which provides smooth reconstructed images while preserving the structure of image edges. An important aspect of this prior is that all of its parameters are automatically estimated from the image data through the EM algorithm. Finally, methods for further improving the accuracy of the method by replacing the one-step-late EM algorithm for updating the image using eq. (7) by another optimization method is an issue of current research.

5. REFERENCES

[1] M. N. Wernick and J. N. Aarsvold, *Emission Tomography: The Fundamentals of PET and SPECT*, Elsevier Science, USA, UK, 2004.

[2] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Transactions on Medical Imaging*, vol. 9, no. 1, pp. 84–93, 1990.

[3] D. M. Higdon, J. E. Bowsher, V. E. Johnson, T. G. Turkington, D. R Gilland, and R. J. Jaczszak, "Fully Bayesian estimation of Gibbs hyperparameters for emission computed tomography data," *IEEE Transactions on Medical Imaging*, vol. 16, no. 5, pp. 516–526, 1997.

[4] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, "Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 687–701, 1994.

[5] N. Dey, L. Blanc-Feraud, C. Zimmer, P. Roux, Z. Kam, J. C. Olivo-Marin, and J. Zerubia, "Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution," *Microscopy Research and Technique*, vol. 69, no. 4, pp. 260–266, 2006.

[6] R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Poisson image reconstruction with total variation regularization," in *Proceedings of the* 17^{th} *IEEE International Conference on Image Processing (ICIP)*, 2010, pp. 4177–4180.

[7] I. T. Hsiao, A. Rangarajan, and G. Gindi, "Joint MAP Bayesian tomographic reconstruction with a Gamma-mixture prior," *IEEE Transactions on Medical Imaging*, vol. 11, no. 12, pp. 1466–14775, 2002.

[8] J. H. Chang and J. M. M. Anderson, "Regularized image reconstruction algorithms for positron emission tomography," *IEEE Transactions on Medical Imaging*, vol. 23, no. 9, pp. 1165–1175, 2004.

[9] I. T. Hsiao, A. Rangarajan, and G. Gindi, "A new convex edge-preserving median prior with apllications to tomography," *IEEE Transactions on Medical Imaging*, vol. 22, no. 5, pp. 580–585, 2003.

[10] Y. Chen, J. Ma, Q. Feng, L. Luo, P. Shi, and W. Chen, "Nonlocal prior Bayesian tomographic reconstruction," *Journal of Mathematical Imaging and Vision*, vol. 30, pp. 133–146, 2008.

[11] H. Ayasso and A. Mohammad-Djafari, "Joint NDT image restoration and segmentation using Gauss-Markov-Potts prior models and variational Bayesian computation," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2265–2277, 2010.

[12] H. Ayasso, S. Fekih-Salem, and A. Mohammad-Djafari, "Variational Bayes approach for tomographic reconstruction," in *Proceedings of the* 28^{th} *International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt)*, Sao Paolo, Brazil, 2008, pp. 243–251.

[13] G. Sfikas, C. Nikou, N. Galatsanos, and C. Heinrich, "Spatially varying mixtures incorporating line processes for image segmentation," *Journal of Mathematical Imaging and Vision*, vol. 36, no. 2, pp. 91–110, 2009.

[14] S. Sanjay-Gopal and T. Hebert, "Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm," *IEEE Transactions on Image Processing*, vol. 7, no. 7, pp. 1014–1028, 1998.

[15] C. Nikou, N. Galatsanos, and A. Likas, "A class-adaptive spatially variant mixture model for image segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1121–1130, 2007.

[16] C. Nikou, A. Likas, and N. Galatsanos, "A Bayesian framework for image segmentation with spatially varying mixtures," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2278–2289, 2010.

[17] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[18] S. Kotz, N. Balakrishnan, and N.L. Johnson, *Continuous multivariate distributions*, Wiley, New York, 2000.

[19] H. R. Sheikh Z. Wang, A. C. Bovik and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.