Segmentation of skin lesions in 2D and 3D ultrasound images using a spatially coherent generalized Rayleigh mixture model

Marcelo Alejandro Pereyra, Nicolas Dobigeon, Hadj Batatia, Jean-Yves Tourneret

Segmentation of skin lesions in 2D and 3D ultrasound images using a spatially coherent generalized Rayleigh mixture model

Marcelo Pereyra, Nicolas Dobigeon, Hadj Batatia and Jean-Yves Tourneret
E-mail: {Marcelo.Pereyra, Nicolas.Dobigeon, Hadj.Batatia, Jean-Yves.Tourneret}@enseeiht.fr

TECHNICAL REPORT – 2011, December
University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse cedex 7, France

Abstract

This paper addresses the problem of jointly estimating the statistical distribution and segmenting lesions in multiple-tissue high-frequency skin ultrasound images. The distribution of multiple-tissue images is modeled as a spatially coherent finite mixture of heavy-tailed Rayleigh distributions. Spatial coherence inherent to biological tissues is modeled by enforcing local dependence between the mixture components. An original Bayesian algorithm combined with a Markov chain Monte Carlo method is then proposed to jointly estimate the mixture parameters and a label-vector associating each voxel to a tissue. More precisely, a hybrid Metropolis-within-Gibbs sampler is used to draw samples that are asymptotically distributed according to the posterior distribution of the Bayesian model. The Bayesian estimators of the model parameters are then computed from the generated samples. Simulation results are conducted on synthetic data to illustrate the performance of the proposed estimation strategy. The method is then successfully applied to the segmentation of in-vivo skin tumors in high frequency 2D and 3D ultrasound images.

Index Terms

Heavy-tailed Rayleigh distribution, mixture model, Potts-Markov field, Bayesian estimation, Gibbs sampler.

I. Introduction

Ultrasound imaging is a longstanding medical imaging modality with important applications in diagnosis, preventive examinations, therapy and image-guided surgery. In dermatologic oncology,

diagnosis relies mainly on surface indicators such as color, shape and texture whereas the two more reliable measures are the depth of the lesion and the number of skin layers that have been invaded. Currently, these can only be evaluated after excision. Recent advances in high frequency transducers and 3D probes have opened new opportunities to perform non-invasive diagnostics using ultrasound images. However, changing dermatological practices requires developing robust segmentation algorithms. Despite the extensive literature on the subject, accurate segmentation of ultrasound images is still a challenging task and a focus of considerable research efforts. Current segmentation techniques are extremely application-specific, developed mainly for echocardiography followed by transrectal prostate examination (TRUS), kidney, breast cancer and (intra) vascular diseases (IVUS) [1]. A survey of the state-of-the-art methods up to 2006 is presented in [1].

Segmentation in echocardiography, TRUS and IVUS is mainly concerned with the detection and tracking of organ boundaries. Lesion delimitation is significantly different and more challenging. On one hand, unlike organs, lesions exhibit soft or "fuzzy" edges that are difficult to capture with boundary detection techniques. On the other, their echogenic and statistical characteristics are visibly different from those of their surrounding tissues. This fact has motivated the development of region-based segmentation techniques as opposed to boundary-based methods, which are still an active research subject in other medical ultrasound domains [2]–[4]. Similarly, lesions do not have anatomically predefined shapes as is the case for organs and are unlikely to benefit in the near future from recent works on anatomical or learned statistical shape priors [5]–[7]. This might change with the improvement of geometric tumor growth models derived from computational biology [8]. Early lesion segmentation methods have focused mainly on thresholding [9], [10] and were superseded by texture-based techniques. Madabhushi *et al.* derived an active contour based on texture and boundary features [11]. Huang *et al.* proposed a texture segmentation technique based on a neural network and a watershed algorithm [12]. In addition, Gaussian mixture models coupled with Markov random fields were proposed to segment lesions based on their region statistics [13], [14]. Moreover, since the seminal work of Dias *et al.* [15], Rayleigh mixtures have become a powerful model for region-based ultrasound image segmentation. The use of Rayleigh instead of Gaussian distributions is strongly justified by the physics of the image formation process that generates B-mode ultrasound images [16]. Based on the assumption that each biological tissue has its proper Rayleigh statistics, tissue

segmentation is achieved by separating the mixture components. This is achieved by finding the maximum-likelihood (ML) or maximum-a-posteriori (MAP) estimators of the lesion contours. The optimization problem stemming from the ML and MAP estimators was solved in [15] using an interactive dynamic programming (IDP) algorithm that jointly estimated the MAP contour and the mixture parameters. The authors performed several experiments on real echocardiography images and showed that the proposed method accurately segments heart walls.

With the development of deformable models, Brusseau *et al.* proposed a statistical parametric active contour (AC) [17]. A parametric AC is a regularized curve defined by a set of points in the image domain that can be moved to maximize the segmentation posterior [18]. In the work of Brusseau *et al.*, the 2-mixture components were separated using a statistical region AC which iteratively estimated the Rayleigh parameter of each component and evolved to optimize the segmentation. Also, given that convergence to a global optimum is not guaranteed, the authors proposed an *ad-hoc* automatic initialization technique. This method was further improved by Cardinal *et al.* [19] who substituted the parametric AC by an edge-based level set (LS) derived from the original work of Osher and Sethian [20]. A second modification was the introduction of an expectation-maximization (EM) algorithm to estimate the mixture parameters during initialization, thus removing the need to estimating them iteratively. The authors reported that the Rayleigh mixture LS method outperforms classic gradient-based level set at intravascular image segmentation. In addition, Saroul *et al.* recently applied the Rayleigh mixture model to prostate segmentation in transrectal ultrasound images [21]. In this case, the LS was replaced by a deformable model based on a super ellipse whose evolution was computed using a variational algorithm. The authors showed that the regularization introduced by this deformable model could compensate partial occlusion.

Rayleigh-mixture models were extended to tissues with generalized Rayleigh statistics by Destrempes *et al.* [22], who proposed a carotid artery segmentation method based on a Nakagami mixture and a deformable model. As in [19], the estimation of the mixture parameters was achieved using an EM algorithm under the assumption that observations are independent. The evolution of the deformable model was computed using exploration/selection, a stochastic optimization algorithm that converges to the global optimum. However, since the mixture parameters are estimated with an EM algorithm, overall global convergence is not guaranteed. One other important contribution is the Rayleigh region-based LS method presented in [23], that adapted

the fundamental work of Chan and Vese [24] on ACs without edges to ultrasound images with Rayleigh statistics. These region-based LS should be very appropriate for ultrasound images of lesions as they are able to segment objects with smooth edges under poor signal-to-noise ratio conditions. This work was recently generalized to all the distributions from the exponential family (i.e., Gamma, Rayleigh, Poisson, etc.) in [25]. However, these methods have not yet been applied to lesion segmentation in ultrasound images.

This paper addresses the problem of jointly estimating the statistical distribution and segmenting lesions in multiple-tissue 2D and 3D high-frequency skin ultrasound images. To our knowledge this is the first ultrasound image segmentation method specific to skin lesions. We propose to model multiple-tissue images using a heavy-tailed Rayleigh mixture, a model that has been inspired by the single-tissue model studied in [26]. The proposed mixture model is equipped with a Markov random field (MRF) that takes into account the spatial correlation inherent to biological tissues. Note that Potts Markov fields are particularly well suited for label-based segmentation as explained in [27] and further studied in [28]–[31]. Potts Markov models enhance segmentation because of their ability to capture the spatial correlation that exists between neighbor class labels [28]. This correlation arises naturally from the spatial organization of biological tissues and is particularly important in skin because of its layered structure. Finally, while the Potts prior is an effective means to introduce spatial correlation between the class labels, it is interesting to mention that other more complex models could have been used instead. In particular, Marroquin *et al.* [32] have shown that better segmentation results may be obtained by using a two-layer hidden field, where hidden labels are assumed to be independent and correlation is introduced at a deeper layer by a vectorial Markov field. Similarly, Woolrich *et al.* [33] have proposed to approximate the Potts field by modeling mixture weights with a Gauss-Markov random field. However, these alternative models are not well adapted for 3D images because they require significantly more computation and memory resources than the Potts model. These overheads result from the fact that they introduce $(K+1)N$ and KN hidden variables respectively, against only N for the Potts model (N being the number of voxels and K the number of classes). In addition, the segmentation problem is solved using a stochastic optimization algorithm with guaranteed global convergence, removing the need for an initial contour or supervised training. The paper is organized as follows: The statistical model used for a voxel of an ultrasound image is introduced in Section II. Section III introduces the Bayesian

model used for the segmentation of ultrasound images. An hybrid Gibbs sampler generating samples asymptotically distributed according to the posterior distribution of this Bayesian model is described in Section IV. Experiments on synthetic and real data are presented in Section V. Conclusions are finally reported in Section VI.

II. Problem Statement

This section describes the mixture model used for ultrasound image voxels[1]. Let $r_n \in \mathbb{R}^+$ denote an observation, or voxel, in an envelope (B-mode) ultrasound image $\boldsymbol{r} = (r_1, \ldots, r_N)^T$ without logarithmic compression. We assume that r_n is defined by means of the widely accepted point scattering model [36]

$$r_n = \left| \sum_{i=1}^{M} a_i \left[p(t_n - \tau_i) + \jmath \tilde{p}(t_n - \tau_i) \right] \right| \tag{1}$$

where M is the total number of punctual scatterers, $p(t) + \jmath\tilde{p}(t)$ denotes the analytic extension of the interrogating pulse $p(t)$, $a_i \in (0, 1)$ is the cross-section of the ith scatterer, $\tau_i \in \mathbb{R}^+$ is the time of arrival of the ith backscattered wave and t_n is the sampling time associated with r_n. Recent works on scattering in biological tissues have established that r_n, as defined above, converges in distribution towards an α-Rayleigh distribution as M increases [26]

$$r_n \underset{M\to\infty}{\overset{d}{\to}} \alpha\mathcal{R}(\alpha_n, \gamma_n) \tag{2}$$

where $\underset{M\to\infty}{\overset{d}{\to}}$ denotes convergence in distribution, the parameters $\alpha_n \in (0, 2]$ and $\gamma_n \in \mathbb{R}^+$ are the characteristic index and spread associated with the nth voxel.

This paper considers the case where the ultrasound image $\boldsymbol{r}$ is made up by multiple biological tissues with high scatter density (i.e., $M \to \infty$), each with its own echogenicity and therefore its proper speckle statistics. In view of this spatial configuration, we propose to model $\boldsymbol{r}$ by an α-Rayleigh stationary process with piecewise constant parameters. More precisely, we assume that there is a set of stationary classes $\{C_1, \ldots, C_K\}$ such that

$$\forall r_n \in \mathcal{C}_k, \quad r_n \sim \alpha\mathcal{R}(\alpha_k, \gamma_k) \tag{3}$$

[1]Part of this work has been presented at IEEE ICASSP'11, Prague, Czech Republic, May 2011 [34], and at EUSIPCO'11, Barcelona, Spain, Sep. 2011 [35].

where α_k and γ_k are the parameters associated with the class $\mathcal{C}_k$ (i.e., the kth biological tissue). As a consequence, it is possible to express the distribution of r_n by means of the following mixture of α-Rayleigh distributions

$$r_n \sim \sum_{k=1}^{K} \omega_k \alpha\mathcal{R}(\alpha_k, \gamma_k) \tag{4}$$

where K is the number of classes and ω_k represents the relative weight (or proportion) of the kth class with $\sum_k \omega_k = 1$. Lastly, to take into account the spatial coherence inherent to biological tissues we will consider that the class of a given voxel depends on those of its neighbors.

It should be noted that the proposed α-Rayleigh mixture model is closely related to two other mixture models. On the one hand it generalizes the Rayleigh mixture model, which has been extensively applied to ultrasound image modeling. On the other, it can be shown that before being transformed by acquisition and demodulation, radio frequency ultrasound signals are distributed according to a symmetric α-stable distribution [26]. Hence, the proposed α-Rayleigh mixture model can be interpreted as a transformation of the symmetric α-stable mixture model studied in [37]. In addition, it is interesting to mention that the α-Rayleigh distribution has been used successfully for SAR images in [38], [39]. The methods proposed in [38], [39] have been recently applied to characterize tissues in annotated ultrasound images [26]. This paper extends those methods by including in the estimation problem the identification of regions in the image with similar α-Rayleigh parameters (each region being associated with a different tissue). This is achieved by proposing a novel Bayesian estimation algorithm based on the α-Rayleigh mixture model (4) coupled with a Markov random field prior that captures the spatial coherence inherent to biological tissues. Finally, akin to [17], [19], [22], [23], note that the model (4) uses a simplified image representation based on regions and does not describe the boundaries between tissues explicitly.

The following section addresses the problem of estimating the parameters of the spatially coherent α-Rayleigh mixture model introduced in (4) and performing the segmentation of ultrasound images.

III. Bayesian Model

A label vector $\boldsymbol{z} = (z_1, \ldots, z_N)^T$ is introduced to map observations $\boldsymbol{r}$ to classes $C_1, \ldots, C_K$ (i.e., $z_n = k$ if and only if $r_n \in C_k$). This label vector will allow each image observation to

be characterized and different kinds of tissues to be discriminated. Note that the weights ω_k are directly related to the labels through the probabilities $\mathrm{P}[z_n = k] = w_k$ for $k = 1, \ldots, K$. Consequently, the unknown parameter vector for the mixture (4) can be defined as $(\boldsymbol{\theta}, \boldsymbol{z})$ where $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\gamma}^T)^T$ with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^T$ and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)^T$. This section studies a Bayesian model associated with $(\boldsymbol{\theta}, \boldsymbol{z})$. This model requires defining the likelihood and the priors for the unknown parameters.

A. Likelihood

Assuming that the observations r_n are independent and using the mixture model (4), the likelihood of the proposed Bayesian model can be written as

$$p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z}) = \prod_{k=1}^{K} \prod_{\{n|z_n=k\}} p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k) \tag{5}$$

where $\{n|z_n = k\}$ denotes the subset of indexes $n = 1, \ldots, N$ that verify $z_n = k$,

$$p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k) = r_n \int_0^\infty \lambda \exp\left[-(\gamma_k\lambda)^{\alpha_k}\right] J_0(r_n\lambda)\, d\lambda \tag{6}$$

is the probability density function (pdf) of an α-Rayleigh distribution with parameters α_k and γ_k and J_0 is the zeroth order Bessel function of the first kind.

B. Parameter priors

1) Labels: It is natural to consider that there is some correlation between the probabilities $\mathrm{P}[z_n = k]$ of a given voxel and those of its neighbors. Since the seminal work of Geman [40], MRFs have become very popular to model neighbor correlation in images. MRFs assume that the distribution of a pixel conditionally to all other pixels of the image equals the distribution of this pixel conditionally to its neighbors. Consequently, it is important to properly define the neighborhood structure. The neighborhood relation between two pixels (or voxels) i and j has to be symmetric: if i is a neighbor of j then j is also a neighbor of i. There are several neighborhood structures that have been used in the literature. In the bidimensional case, neighborhoods defined by the four or eight nearest voxels represented in Fig. 1 are the most commonly used. Similarly, in the tridimensional case the most frequently used neighborhoods are defined by the six or fourteen nearest voxels represented in Fig 2. In the rest of this paper 4-pixel neighborhoods will be considered for 2D images and 6-voxel neighborhoods for 3D images. Therefore, the associated

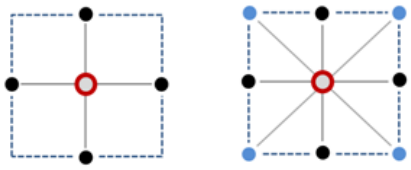


Fig. 1. 4-pixel (left) and 8-pixel (right) neighborhood structures. The pixel considered appears as a void red circle whereas its neighbors are depicted in full black and blue.

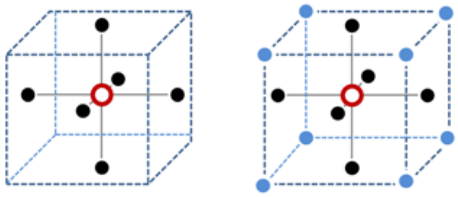


Fig. 2. 6-voxel (left) and 14-voxel (right) neighborhood structures. The voxel considered appears as a void red circle whereas its neighbors are depicted in full black and blue.

set of neighbors, or cliques, can only have vertical, horizontal and depth configurations (see [40], [41] for more details).

Once the neighborhood structure has been established, the MRF can be defined. Let z_n denote the random variable indicating the class of the nth image voxel. In the case of K classes, the random variables $z_1, \ldots, z_N$ take their values in the finite set $\{1, \ldots, K\}$. The whole set of random variables $\boldsymbol{z}$ forms a random field. An MRF is then defined when the conditional distribution of z_n given the other pixels $\boldsymbol{z}_{-n} = (z_1, \ldots, z_{n-1}, z_{n+1}, \ldots, z_N)$ only depends on its neighbors $\boldsymbol{z}_{\mathcal{V}(n)}$, i.e.,

$$\mathrm{P}[z_n|\boldsymbol{z}_{-n}] = \mathrm{P}[z_n|\boldsymbol{z}_{\mathcal{V}(n)}] \tag{7}$$

where $\mathcal{V}(n)$ contains the neighbors of z_n according to the neighborhood structure considered.

In this study we will first consider 2D and 3D Potts Markov fields as prior distributions for $\boldsymbol{z}$. More precisely, 2D MRFs are considered for single-slice (2D) ultrasound images whereas 3D MRFs are used for multiple-slice (3D) images. In light of the Hammersley-Clifford theorem, the corresponding prior for $\boldsymbol{z}$ can be expressed as follows:

$$p(\boldsymbol{z}) = \frac{1}{C(\beta)} \exp\left[\sum_{n=1}^{N} \sum_{n' \in \mathcal{V}(n)} \beta \delta(z_n - z_{n'})\right] \tag{8}$$

where β is the granularity coefficient, $C(\beta)$ is the normalizing constant or partition function [42] and $\delta(\cdot)$ is the Kronecker function. The hyperparameter β tunes the degree of homogeneity of each region in the image. A small value of β induces a noisy image with a large number of regions, contrary to a large value of β that leads to few and large homogeneous regions. In this work, the granularity coefficient β will be fixed a priori. However, it is interesting to mention that the estimation of β has been receiving a lot of attention in the literature [31], [43]–[46]. Estimating the granularity coefficient using one of these methods is clearly an interesting problem that will be investigated in future work. Finally, it is interesting to note that despite not knowing $C(\beta)$, drawing labels $\boldsymbol{z} = (z_1, \ldots, z_N)$ from the distribution (8) can be easily achieved by using a Gibbs sampler [47].

2) α-Rayleigh parameters: The prior for each characteristic index α_k ($k = 1, \ldots, K$) is a uniform distribution on $(0, 2]$

$$\alpha_k \sim \mathcal{U}(0, 2). \tag{9}$$

This choice is motivated by the fact that the only information available a priori about this parameter, is that it can take values in the interval $(0, 2]$.

The prior for each spread parameter γ_k is an inverse gamma distribution with hyperparameters a_0 and b_0

$$\gamma_k \sim \mathcal{IG}(a_0, b_0), \quad k = 1, \ldots, K. \tag{10}$$

This choice is motivated by the fact that the inverse gamma distribution allows either very vague or more specific prior information to be incorporated depending on the choice of the hyperparameters a_0 and b_0 ($a_0 = b_0 = 1$ will be used in our experiments corresponding to a vague prior distribution).

Assuming a priori independence between the parameters α_k and γ_k, the prior for $\boldsymbol{\theta}$ is

$$p(\boldsymbol{\theta}) = p(\boldsymbol{\alpha})p(\boldsymbol{\gamma}) = \prod_{k=1}^{K} p(\alpha_k)p(\gamma_k). \tag{11}$$

We will also assume that the α-Rayleigh parameters are independent from the labels associated with the image voxels. Thus the joint prior for the unknown parameters $(\boldsymbol{\theta}, \boldsymbol{z})$ can be expressed as

$$p(\boldsymbol{\theta}, \boldsymbol{z}) = p(\boldsymbol{z})\, p(\boldsymbol{\theta}) \tag{12}$$

where $p(\boldsymbol{z})$ has been defined in (8) and $p(\boldsymbol{\theta})$ in (11).

Figure 3 presents the proposed Bayesian model as a directed acyclic graph (DAG) summarizing the relationships between the different parameters and hyperparameters.

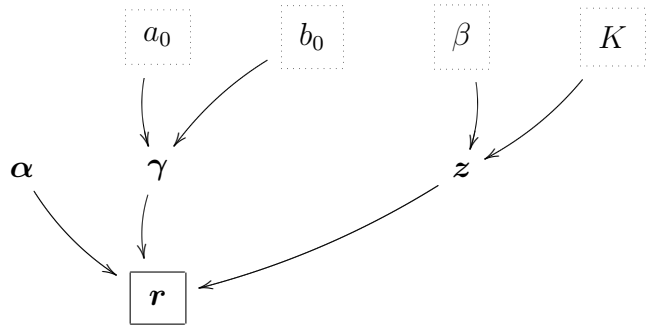


Fig. 3. Directed acyclic graph (DAG) for the α-Rayleigh mixture model (the fixed nonrandom hyperparameters appear in dashed boxes).

C. Posterior Distribution of $(\boldsymbol{\theta}, \boldsymbol{z})$

Using Bayes theorem, the posterior distribution of $(\boldsymbol{\theta}, \boldsymbol{z}) = (\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{z})$ can be expressed as follows

$$\begin{aligned} p\left(\boldsymbol{\theta}, \boldsymbol{z}|\boldsymbol{r}\right) &= \frac{p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z})p(\boldsymbol{\theta}, \boldsymbol{z})}{p(\boldsymbol{r})} \\ &\propto p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z})p(\boldsymbol{\theta}, \boldsymbol{z}) \end{aligned} \tag{13}$$

where $\propto$ means "proportional to" and the likelihood $p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z})$ and the joint prior $p(\boldsymbol{\theta}, \boldsymbol{z})$ have been defined in (5) and (12).

Unfortunately the posterior distribution (13) is too complex to derive closed form expressions for the minimum mean square error (MMSE) or MAP estimators of the unknown parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$ and $\boldsymbol{z}$ [2]. One can think of using the EM algorithm [48] that has received much attention for mixture problems (see [19], [22] for applications to ultrasound images). However, EM algorithms have many known shortcomings. For instance, they suffer from *convergence to local maxima or saddle points of the log-likelihood function and sensitivity to starting values* [49, p. 259]. Note that analyzing the concavity properties of the logarithm of (5) is not easy because the α-Rayleigh distribution does not belong to the exponential family. An interesting alternative is to

[2]note that $p(\boldsymbol{\theta}, \boldsymbol{z})$ involves the potential of a Potts Markov field and its intractable partition function $C(\beta)$ and that $p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z})$ is the product of N indefinite integrals

use a Markov Chain Monte Carlo (MCMC) method generating samples that are asymptotically distributed according to the target distribution (13) [47]. The generated samples are then used to approximate the Bayesian estimators. This strategy has been used successfully in many image processing applications [50]–[54]. One sampling technique allowing the parameters of ultrasound images to be estimated is studied in the next section.

IV. Hybrid Gibbs Sampler

This section studies a hybrid Metropolis-within-Gibbs sampler for generating samples that are asymptotically distributed according to (13). The histogram of the generated samples is guaranteed to converge to the posterior (13) [47, p. 269]. One of the most popular methods for generating samples distributed according to a distribution whose pdf or probability masses are known up to a multiplicative constant is the Gibbs sampler. The conventional Gibbs sampler draws samples according to the conditional distributions associated with the distribution of interest (here the posterior (13)). When a conditional distribution cannot be sampled easily, one can resort to a Metropolis-Hastings (MH) move, which generates samples according to an appropriate proposal and accept or reject these generated samples with a given probability. The resulting sampler is referred to as Metropolis-within-Gibbs sampler (see [47] for more details about MCMC methods). The sampler investigated in this section is based on the conditional distributions $\mathrm{P}[\boldsymbol{z}|\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{r}]$, $p(\boldsymbol{\alpha}|\boldsymbol{z},\boldsymbol{\gamma},\boldsymbol{r})$ and $p(\boldsymbol{\gamma}|\boldsymbol{z},\boldsymbol{\alpha},\boldsymbol{r})$ that are described in the next paragraphs (see also Algorithm 1 below).

A. *Conditional probability* $\mathrm{P}[\boldsymbol{z}|\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{r}]$

The label vector $\boldsymbol{z}$ can be updated coordinate-by-coordinate using Gibbs moves. More precisely, the conditional probabilities $\mathrm{P}[z_n|\boldsymbol{z}_{-n},r_n,\alpha_k,\gamma_k]$ can be computed using the Bayes rule

$$\mathrm{P}[z_n=k|\boldsymbol{z}_{-n},r_n,\alpha_k,\gamma_k] \propto p(r_n|z_n=k,\boldsymbol{\alpha},\boldsymbol{\gamma})p(z_n|\boldsymbol{z}_{-n}) \tag{14}$$

where $k=1,\ldots,K$ (it is recalled that K is the number of classes) and where $\boldsymbol{z}_{-n}$ is the vector $\boldsymbol{z}$ whose nth element has been removed. These posterior probabilities can be expressed as

$$\begin{aligned}\mathrm{P}[z_n=k|\boldsymbol{z}_{-n},r_n,\alpha_k,\gamma_k] \propto \pi_{n,k} \triangleq \exp\left[\sum_{n'\in\mathcal{V}(n)}\beta\delta(k-z_{n'})\right] \\ \times\, r_n\int_0^\infty \lambda\exp\left[-(\gamma_k\lambda)^{\alpha_k}\right]J_0(r_n\lambda)\,d\lambda.\end{aligned} \tag{15}$$

Algorithm 1 Proposed Hybrid Gibbs Sampler

Initialization:

– Sample α_k^0 $(k = \{1, \ldots, K\})$ from the pdf in (9).

– Sample γ_k^0 $(k = \{1, \ldots, K\})$ from the pdf in (10).

– Generate $z_1^0, z_2^0, \ldots, z_N^0$ with probabilities $P[z_n^0 = k] = \frac{1}{K}$.

for $t = 1, 2, \ldots$ to T **do**

— *Update* $\boldsymbol{\alpha}$ —

for $k = 1, 2, \ldots$ to K **do**

1. Propose $\alpha_k^* \sim \mathcal{N}_{(0,2)}(\alpha_k^{(t-1)}, \sigma_{\alpha,k}^2)$ (see (17)).

2. Compute the acceptance ratio using expression (18).

3. Draw $u \sim \mathcal{U}(0, 1)$.

if $(u <$ ratio$)$ **then**

4. Set $\alpha_k^{(t)} = \alpha_k^*$.

else

5. Set $\alpha_k^{(t)} = \alpha_k^{(t-1)}$.

end if

end for

— *Update* $\boldsymbol{\gamma}$ —

for $k = 1, 2, \ldots$ to K **do**

6. Propose $\gamma_k^* \sim \mathcal{N}^+(\gamma_k^{(t-1)}, \sigma_{\gamma,k}^2)$ (see (19)).

7. Compute the acceptance ratio using expression (20).

8. Draw $u \sim \mathcal{U}(0, 1)$.

if $(u <$ ratio$)$ **then**

9. Set $\gamma_k^{(t)} = \gamma_k^*$.

else

10. Set $\gamma_k^{(t)} = \gamma_k^{(t-1)}$.

end if

end for

— *Update* $\boldsymbol{z}$ —

for $n = 1, 2, \ldots$ to N **do**

11. Draw z_n from $\{1, \ldots, K\}$ with probabilities (16).

end for

end for

The integral $r_n \int_0^\infty \lambda \exp\left[-(\gamma_k \lambda)^{\alpha_k}\right] J_0(r_n \lambda)\, d\lambda$ is evaluated using the approximations presented in paragraph IV-D. Once all the quantities $\pi_{n,k}$, $k = 1, \ldots, K$, have been computed, they are normalized to obtain the posterior probabilities $\tilde{\pi}_{n,k} \triangleq \mathrm{P}[z_n = k | \boldsymbol{z}_{-n}, r_n, \alpha_k, \gamma_k]$ as follows

$$\tilde{\pi}_{n,k} = \frac{\pi_{n,k}}{\sum_{k=1}^{K} \pi_{n,k}}. \tag{16}$$

Note that the posterior probabilities of the label vector $\boldsymbol{z}$ in (15) and (16) define an MRF. Finally, samples z_n are generated by drawing discrete variables from $\{1, \ldots, K\}$ with the respective probabilities $\{\tilde{\pi}_{n,1}, \ldots, \tilde{\pi}_{n,K}\}$. Because of its large dimension, sampling $\boldsymbol{z}$ according to (16) is the most computationally intensive step of the proposed hybrid Gibbs sampler. Therefore it is important to chose an efficient implementation for this step. In this work $\boldsymbol{z}$ has been sampled using a parallel chromatic Gibbs sampler [55].

B. Conditional probability density function $p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{r})$

The conditional pdf $p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{r})$ can be expressed as follows

$$p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{r}) \propto p(\boldsymbol{r}|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{z}) p(\boldsymbol{\alpha})$$

where $p(\boldsymbol{r}|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{z})$ is defined in (5) and $p(\boldsymbol{\alpha}) = \prod_{k=1}^{K} p(\alpha_k)$. The generation of samples according to $p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{r})$ is not easy to perform. We propose in this paper to sample $\boldsymbol{\alpha}$ coordinate-by-coordinate using MH moves. In this work, the proposal distribution is a truncated normal distribution centered on the previous value of the chain with variance $\sigma^2_{\alpha,k}$

$$\alpha_k^* \sim \mathcal{N}_{(0,2)}(\alpha_k^{(t-1)}, \sigma^2_{\alpha,k}) \tag{17}$$

where α_k^* denotes the proposed value at iteration t and $\alpha_k^{(t-1)}$ is the previous state of the chain. The hyperparameters $\sigma^2_{\alpha,k}$ are adjusted during the burn-in period to ensure an acceptance ratio close to $\frac{1}{3}$, as recommended in [56, p. 316]. This adjustment is performed dynamically by a feedback loop that increases or decreases $\sigma^2_{\alpha,k}$ depending on α_k's acceptance ratio over the last 50 iterations. Note that the proposal (17) results from the so-called random walk MH algorithm [47, p. 245]. Finally, since the prior for α_k is uniform, the MH acceptance rate of the proposed move can be expressed as follows

$$\text{ratio} = \min\left\{1, \frac{\mathcal{N}_{(0,2)}(\alpha_k^{(t-1)}|\alpha_k^*, \sigma^2_{\alpha,k})}{\mathcal{N}_{(0,2)}(\alpha_k^*|\alpha_k^{(t-1)}, \sigma^2_{\alpha,k})} \times \prod_{\{n|z_n=k\}}^{N} \frac{p_{\alpha\mathcal{R}}(r_n|\alpha_k^*, \gamma_k)}{p_{\alpha\mathcal{R}}(r_n|\alpha_k^{(t-1)}, \gamma_k)}\right\} \tag{18}$$

where the likelihoods $p_{\alpha\mathcal{R}}(r_n|\alpha_k^*, \gamma_k)$ and $p_{\alpha\mathcal{R}}(r_n|\alpha_k^{(t-1)})$ have been computed using the approximations described in Section IV-D.

C. Conditional probability density function $p(\boldsymbol{\gamma}|\boldsymbol{\alpha}, \boldsymbol{z}, \boldsymbol{r})$

The conditional pdf $p(\boldsymbol{\gamma}|\boldsymbol{\alpha}, \boldsymbol{z}, \boldsymbol{r})$ can be expressed as follows

$$p(\boldsymbol{\gamma}|\boldsymbol{\alpha}, \boldsymbol{z}, \boldsymbol{r}) \propto p(\boldsymbol{r}|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{z})p(\boldsymbol{\gamma})$$

where $p(\boldsymbol{r}|\boldsymbol{\alpha}, \boldsymbol{\gamma}, \boldsymbol{z})$ is defined in (5) and $p(\boldsymbol{\gamma}) = \prod_{k=1}^{K} p(\gamma_k)$. Again, we propose to sample $\boldsymbol{\gamma}$ coordinate-by-coordinate by using MH moves. The proposal distribution associated with this move is a truncated normal distribution centered on the previous value of the chain with variance $\sigma_{\gamma,k}^2$

$$\gamma_k^* \sim \mathcal{N}_{\mathbb{R}^+}\left(\gamma_k^{(t-1)}, \sigma_{\gamma,k}^2\right) \tag{19}$$

where γ_k^* denotes the proposed value at iteration t, $\gamma_k^{(t-1)}$ is the previous state of the chain and $\mathcal{N}_{\mathbb{R}^+}$ is the Gaussian distribution truncated on $\mathbb{R}^+$. The acceptance ratio for this move is

$$\text{ratio} = \min\left\{1, \frac{\mathcal{N}_{\mathbb{R}^+}\left(\gamma_k^{(t-1)}|\gamma_k^*, \sigma_{\gamma,k}^2\right)}{\mathcal{N}_{\mathbb{R}^+}\left(\gamma_k^*|\gamma_k^{(t-1)}, \sigma_{\gamma,k}^2\right)} \times \prod_{\{n|z_n=k\}}^{N} \frac{p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k^*)p(\gamma_k^*|a_0, b_0)}{p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k^{(t-1)})p(\gamma_k^{(t-1)}|a_0, b_0)}\right\} \tag{20}$$

where the prior distribution $p(\gamma_k|a_0, b_0)$ has been defined in (10). Again, the likelihoods $p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k^*)$ and $p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k^{(t-1)})$ have been computed using the approximations described in Section IV-D.

In the particular case $\alpha_k = 2$, the likelihood simplifies to a Rayleigh distribution for which the prior $p(\gamma_k) = \mathcal{IG}(a_0, b_0)$ is conjugate. As a result the generation of samples from the posterior $p(\gamma_k|\alpha_k, \boldsymbol{z}, \boldsymbol{r})$ reduces to drawing samples from the following inverse gamma distribution

$$\gamma_k^{(t)} \sim \mathcal{IG}\left(a_0 + \sum_{\{n|z_n=k\}}^{N} 1, b_0 + \frac{1}{2}\sum_{\{n|z_n=k\}}^{N} r_n^2\right) \tag{21}$$

where we recall that $a_0 = 1$ and $b_0 = 1$.

D. Approximation of the Likelihood

Evaluating the likelihood function defined in (5) involves the computation of the following indefinite integral

$$\int_0^\infty \lambda \exp\left[-(\gamma_k\lambda)^{\alpha_k}\right] J_0(r_n\lambda)\, d\lambda. \tag{22}$$

In the case where observations are represented using 8-bit precision (i.e., 256-gray levels) the integral can be pre-computed for each level and stored in a look-up-table. The data used in this work is represented using 32-bit precision and the integral had to be solved numerically. This computation is time-consuming and is required for every observation and at every step of the sampler. An efficient way to alleviate this computational complexity is to use the following asymptotic expansions [57], [58][57], [58][57], [58]

$$p_{\alpha\mathcal{R}}(r_n|\alpha_k,\gamma_k) = \sum_{p=0}^{P} a_p r_n^{2p+1} + o\left(r_n^{2(P+1)+1}\right) \tag{23}$$

as $r_n \to 0$ and

$$p(r_n|\alpha_k,\gamma_k) = \sum_{p=1}^{P} b_p r_n^{-\alpha_k p-1} + o\left(r_n^{-\alpha_k(P+1)-1}\right) \tag{24}$$

as $r_n \to \infty$, where the coefficients a_p and b_p are

$$a_p = \frac{1}{\alpha_k\gamma_k}\frac{(-1)^p}{(p!)^2 2^{2p}}\Gamma\left(\frac{2p+2}{\alpha_k}\right)\gamma_k^{-2p-1}$$

$$b_p = \frac{(-1)^{p-1}\,2^{p\alpha_k+1}}{p!\,\pi\gamma_k}\Gamma^2\left(\frac{p\alpha_k+2}{2}\right)\sin\left(\frac{p\pi\alpha_k}{2}\right)\gamma_k^{p\alpha_k+1}.$$

The decision between using (23) or (24) for a particular value r_n has been determined by a threshold which has been computed off-line. This threshold and the choice of P have been studied empirically by comparing (23) and (24) to a numerical solution of the true density (5). Appropriate threshold and P values have been selected off-line for different values of α_k and stored in a look-up-table that is used by the proposed algorithm. Other considerations regarding the implementation of (23) and (24) have been studied in [57].

V. Experimental Results

This section presents experimental results conducted on synthetic and real data to assess the performance of the proposed α-Rayleigh mixture model and the associated Bayesian estimation algorithm. In these experiments the algorithm convergence has been assessed using the "between-within variance criterion", initially studied by Gelman and Rubin [59] and often used to monitor convergence [60, p. 33]. This criterion requires running M parallel chains of length L with different starting values and computing the so-called *potential scale reduction factor* (PSRF) that compares the between-sequence and within-sequence variances [59]. A PSRF close to 1

indicates good convergence of the sampler. In our experiments we have observed PSRF values smaller than 1.01 which confirm the good convergence of the sampler (a PSRF bellow 1.2 is recommended in [61, p. 332]). These values were computed using $M = 25$ parallel chains of length $L = 1,000$ whose first 900-steps were discarded.

A. Synthetic Data

To validate the proposed Bayesian method under controlled ground truth conditions (i.e., known true class labels $\boldsymbol{z}$ and statistical parameters $(\boldsymbol{\alpha}, \boldsymbol{\gamma})$), the algorithm described in Section IV was first applied to the synthetic 3-component α-Rayleigh mixture displayed in Fig. 4(a). The parameters associated with the mixture components of the 3 different 2D regions are $\boldsymbol{\alpha} = [1.99, 1.99, 1.8]^T$ and $\boldsymbol{\gamma} = [1, 5, 10]^T$. Figure 4(b) shows the resulting observation vector $\boldsymbol{r}$, which is the only input provided to the algorithm. Note that the different observations are clearly spatially correlated. The proposed Gibbs sampler has been run for this example using a two-dimensional random field with a 4-pixel neighborhood structure and a granularity coefficient $\beta = 1$. Figure 5 shows histograms of the parameters generated by the proposed Gibbs sampler. These histograms are in good agreement with the actual values of the different parameters. Moreover, the MMSE estimates and the corresponding standard deviations for the different parameters are reported in Table I. These estimates have been computed from a single Markov chain of $25,000$ iterations whose first 100 iterations (burn-in period) have been removed. The MMSE estimates are clearly in good agreement with the actual values of the α-Rayleigh mixture components. Figure 4(c) shows the class labels estimated by the MAP rule applied to the last samples of the Markov chain. The three classes are recovered with a few misclassifications due to the complexity of the problem.

In order to illustrate the effect of the granularity parameter, we have considered other values of the parameter β. Fig. 4(d) and (e) show the class labels obtained with $\beta = 1.2$ and $\beta = 0.8$. We observe that increasing β from 1.0 to 1.2 reduces significantly the number of isolated misclassifications at the expense of increasing errors at the boundaries between the different classes. Decreasing β from 1 to 0.8 increases the number of misclassifications both at the boundaries and within regions.

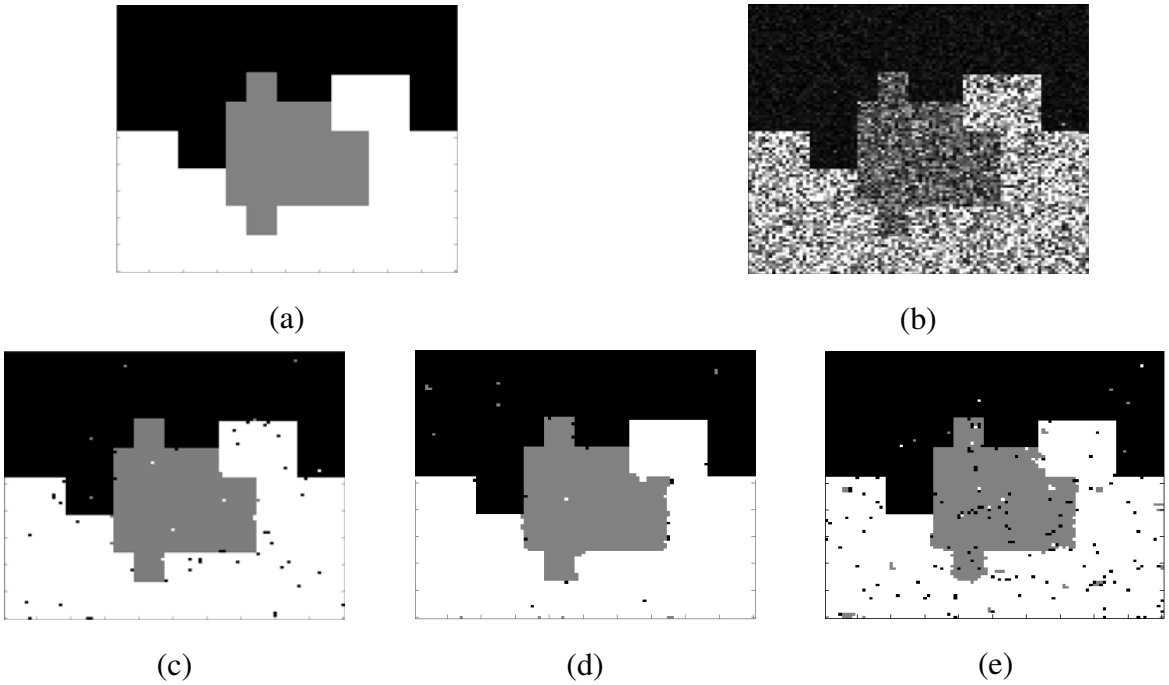


Fig. 4. (a) True labels, (b) observations, MAP label estimates for (c) $\beta = 1$, (d) $\beta = 1.2$ and (e) $\beta = 0.8$.

TABLE I

PARAMETER ESTIMATION

	true value	MMSE estimates	standard deviation
α_1	1.99	1.99	0.002
γ_1	1.00	1.00	0.003
α_2	1.99	1.99	0.003
γ_2	5.00	5.01	0.025
α_3	1.80	1.79	0.006
γ_3	10.00	9.96	0.036

B. Simulated 3D ultrasound image

The synthetic image studied previously is a toy image that differs from a real ultrasound image in many aspects. These aspects include the spatial organization of skin tissue as well as the different physical phenomena intervening in the formation of ultrasound images (i.e., noise, limited spatial resolution, voxel anisotropy, attenuation, etc.). In order to consider a more realistic

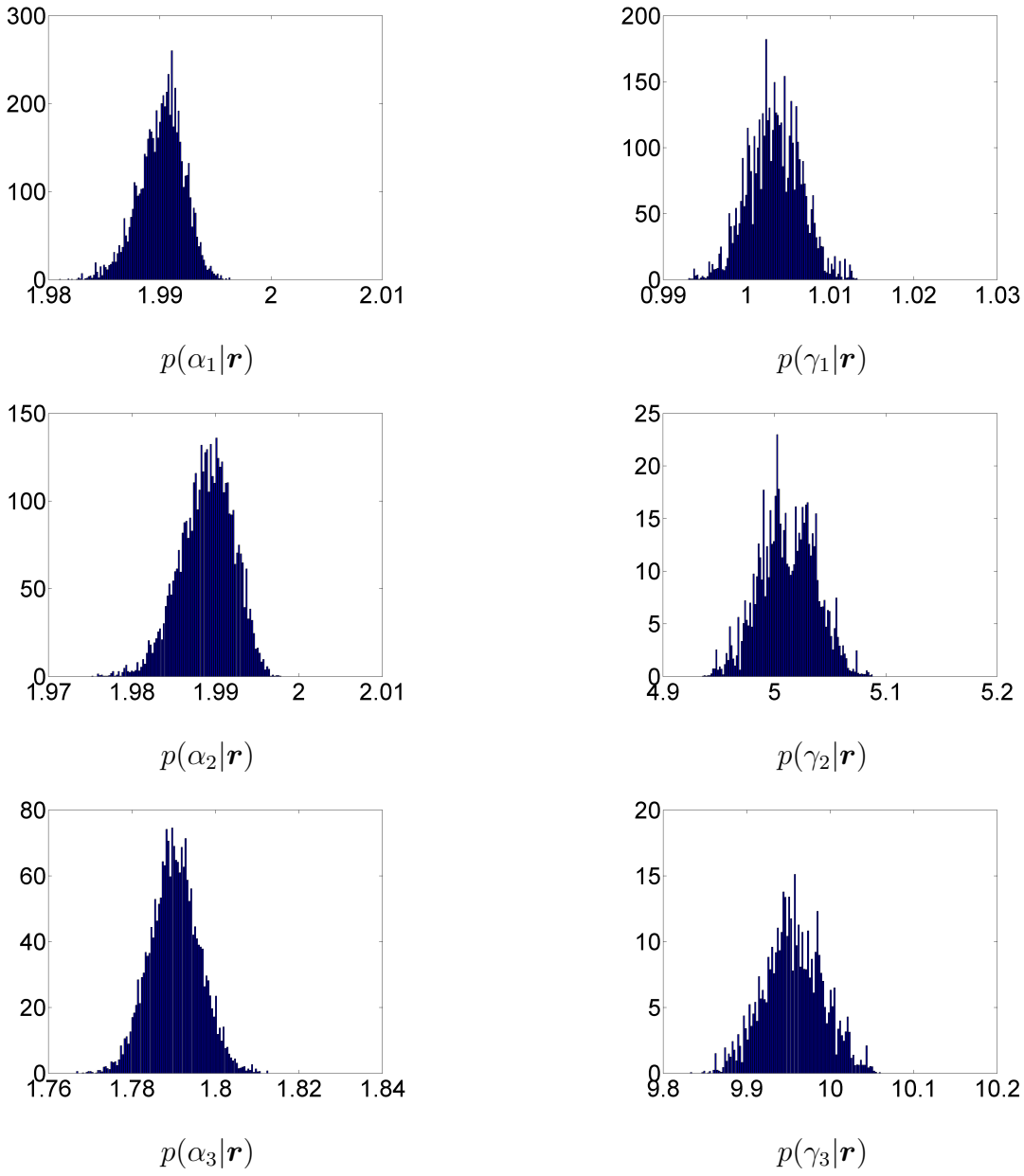


Fig. 5. Histograms of parameters generated using the proposed Gibbs sampler.

scenario, the second set of experiments considers a simulated 3D phantom of skin tissue. This 3D phantom image has been simulated using a 3D ultrasound simulator [62], which has been configured with the parameters of the *dermocup* ultrasound system (Atys Medical, France) used in the in-vivo experiments of section V-C. Three slices of the 30-slice 3D phantom are shown in Figs. 6(a), 6(b) and 6(c). The size of each slice is 400×300 pixels. These images are displayed

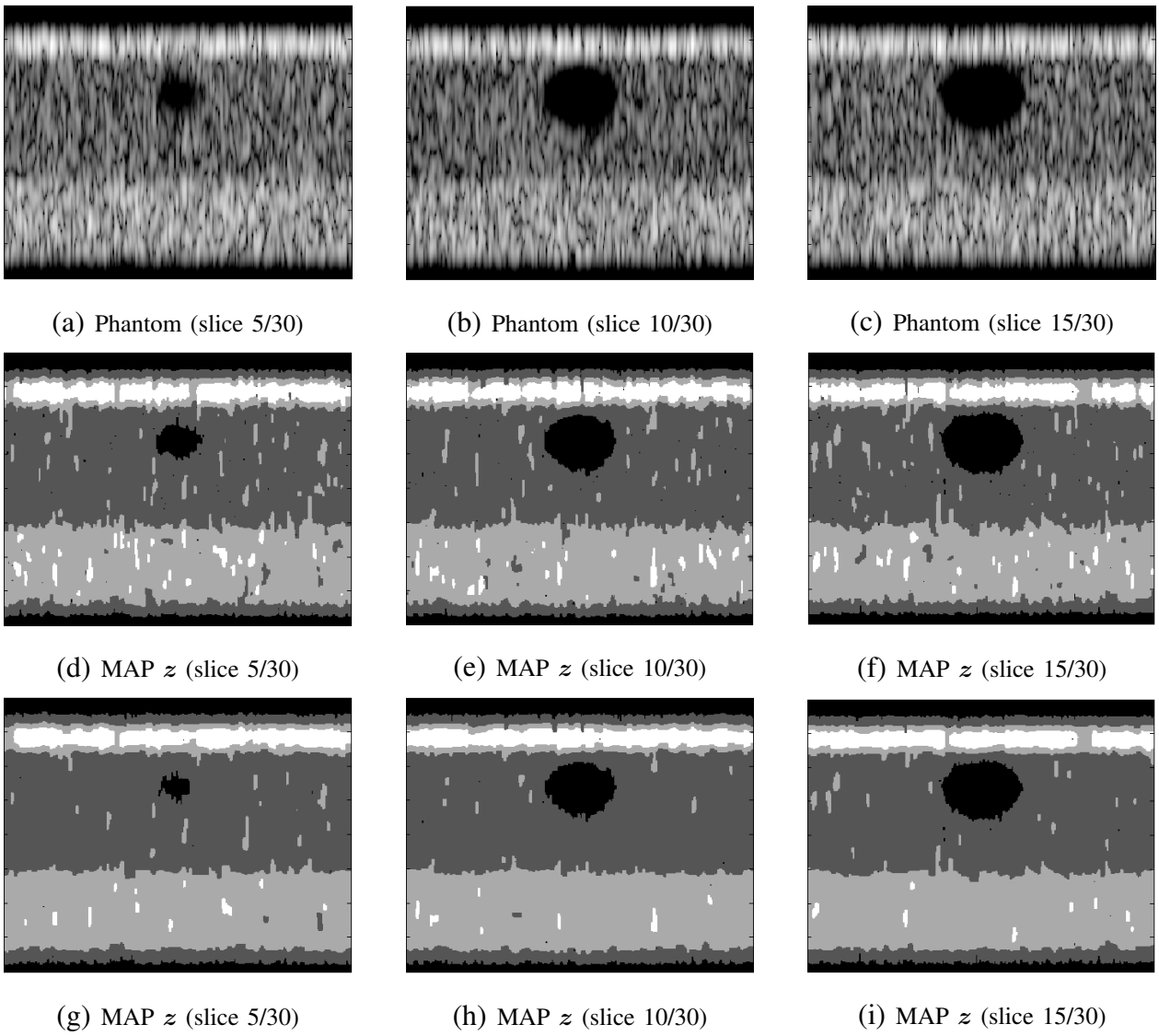
using logarithmic compression; however the proposed algorithm has been applied to B-mode images in linear scale. The 3D skin phantom contains three skin layers (epidermis, papillary dermis and reticular dermis), and one ellipsoidal intra-dermic lesion. Figs. 6(d)-(f) and Figs. 6(g)-(i) show the corresponding MAP estimated labels obtained with the proposed method using $\beta = 1.0$ and $\beta = 1.2$. We observe that in both cases the skin layers and the lesion are clearly recovered with a few misclassifications due to the complexity of the problem. The number of classes for this experiment has been set to $K = 4$ since there are 3 types of healthy tissue in addition to the lesion. These results were computed using a 3D MRF and a single Markov chain of $1,000$ iterations whose first 900 iterations (burn-in period) have been removed.

C. Application to real data

After validating the proposed Gibbs sampler on synthetic data, this section applies the proposed algorithm to the segmentation of two skin lesions. Experiments were conducted using 3D high frequency B-mode ultrasound images of in-vivo skin tissues. These were acquired with a *dermocup* system (Atys Medical, France), equipped with a single-element focalized 25MHz wide-band (40-percent) probe sampled at 100MHz with a $53\mu m$ mechanic lateral step. The proposed α-Rayleigh mixture model describes the statistics of envelope (B-mode) ultrasound images without logarithmic compression [26]. Therefore all experiments have been conducted using this type of data. However, to simplify their visual interpretation, results are displayed using logarithmic compression, which is a standard practice in ultrasound imaging [63]. Note that since α-Rayleigh envelope signals arise from symmetric α-stable radio frequency signals [26] it would be possible to apply the proposed method directly to the radio frequency ultrasound image by replacing the α-Rayleigh mixture model (5) by a symmetric α-stable mixture model [37].

In this work the number of classes K is assumed to be known a-priori. This important parameter is set by the dermatologist who determines visually the number of tissues within the region to be processed. For skin tissues the number of classes depends on the number of layers contained in that region (i.e., epidermis, papillary (upper) dermis, reticular (lower) dermis, hypodermis) in addition to the lesion. More details regarding the number of classes are available in section V-C6.

The Potts granularity coefficient β has been chosen heuristically by testing a few values



(a) Phantom (slice 5/30) (b) Phantom (slice 10/30) (c) Phantom (slice 15/30)

(d) MAP $\boldsymbol{z}$ (slice 5/30) (e) MAP $\boldsymbol{z}$ (slice 10/30) (f) MAP $\boldsymbol{z}$ (slice 15/30)

(g) MAP $\boldsymbol{z}$ (slice 5/30) (h) MAP $\boldsymbol{z}$ (slice 10/30) (i) MAP $\boldsymbol{z}$ (slice 15/30)

Fig. 6. Simulated (log-compressed) US images of skin layers with an intradermic lesion and the corresponding estimated labels. Figs. (a)-(c) depict three slices of the 30-slice 3D digital phantom. MAP label estimates for (d)-(f) $\beta = 1$ and (g)-(i) $\beta = 1.2$.

between 0.5 and 1.5. These tests have suggested that segmentation results best agree with expert annotations for $\beta \in (1, 1.5)$. Finally, β was set to 1 in order to minimize the risk of over-smoothing the segmentation results, which was the main concern of dermatologists. Section V-C4 presents segmentation results obtained with other values of β with discussions. Future work will study the estimation of β jointly with the other unknown parameters of the model, as in [52].

1) Justification of the α-Rayleigh mixture model: the α-Rayleigh mixture model used in this work is based on the assumption that the statistics of single-tissue regions can be well described by an α-Rayleigh distribution. To support this assumption Fig. 7 compares the histogram obtained from a B-mode ultrasound image of in-vivo forearm dermis with the corresponding α-Rayleigh, Nakagami and Gamma distribution fits (additional fits are provided in [26]). To better illustrate fitting at the tails, Fig. 7 displays the probability density functions in logarithmic scale. We observe that the α-Rayleigh distribution provides the best fit and is the only one to accurately describe the heavy-tail of the histogram.

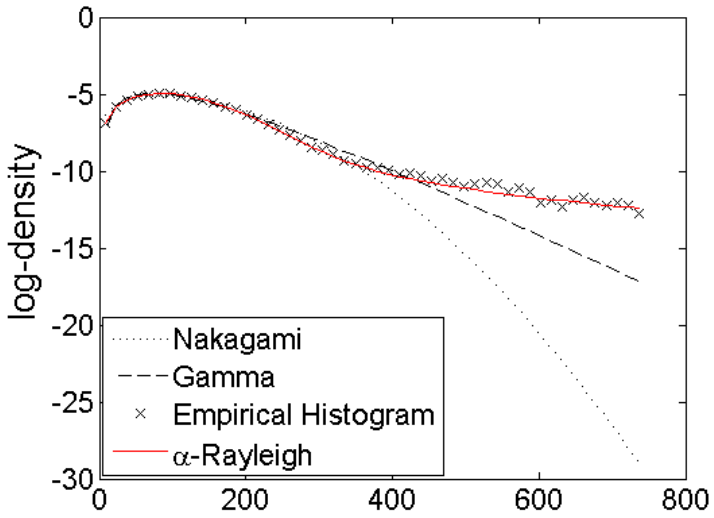
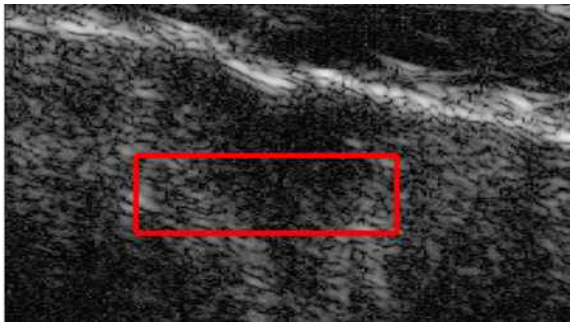


Fig. 7. Comparison of the B-mode histogram obtained from forearm dermis, and the corresponding estimations using the Nakagami, Gamma and α Rayleigh distributions. Plots presented in logarithmic scale to illustrate fitting at the tails.

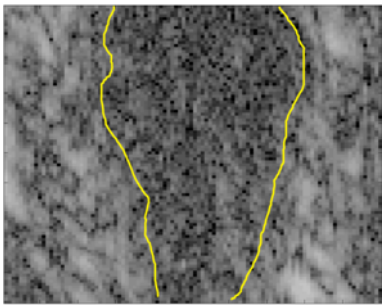
2) Preliminary 2D and 3D experiments: The two following experiments illustrate the importance of introducing spatial correlation between the mixture components. Fig. 8(a) shows a skin lesion outlined by the red rectangle. This region is displayed with coarse expert annotations (yellow curve) in Fig. 8(b). It should be noted that annotations approximately localize the lesion and do not represent an exact ground truth. The following experiments have been conducted with granularity coefficient $\beta = 1$ and the number of classes $K = 2$ since there are only two types of tissue (i.e., lesion and healthy reticular dermis) within the region of interest (ROI). The results have been computed from a single Markov chain of $1,000$ iterations whose first 900 iterations

(burn-in period) have been removed.

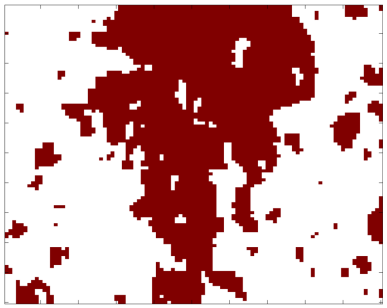
First, the proposed Bayesian algorithm was used to label each voxel of the ultrasound image as *healthy* or *lesion* tissue. The estimated labels obtained using a bidimensional random field are displayed in Fig. 8(c). For comparison purposes, Fig. 8(d) shows the estimation results when labels are considered a priori independent, as in [34]. Due to the proposed MRF prior for the labels, the spatial correlations between image voxels are clearly recovered with the proposed segmentation procedure.



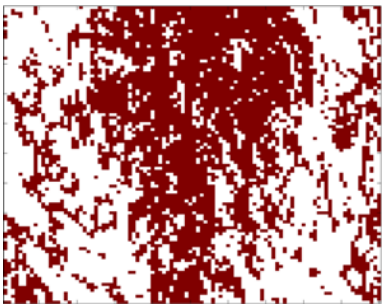
(a) Dermis view with skin lesion (ROI $= 100 \times 100 \times 3$).



(b) ROI (slice 2)



(c) MRF Labels $\boldsymbol{z}$



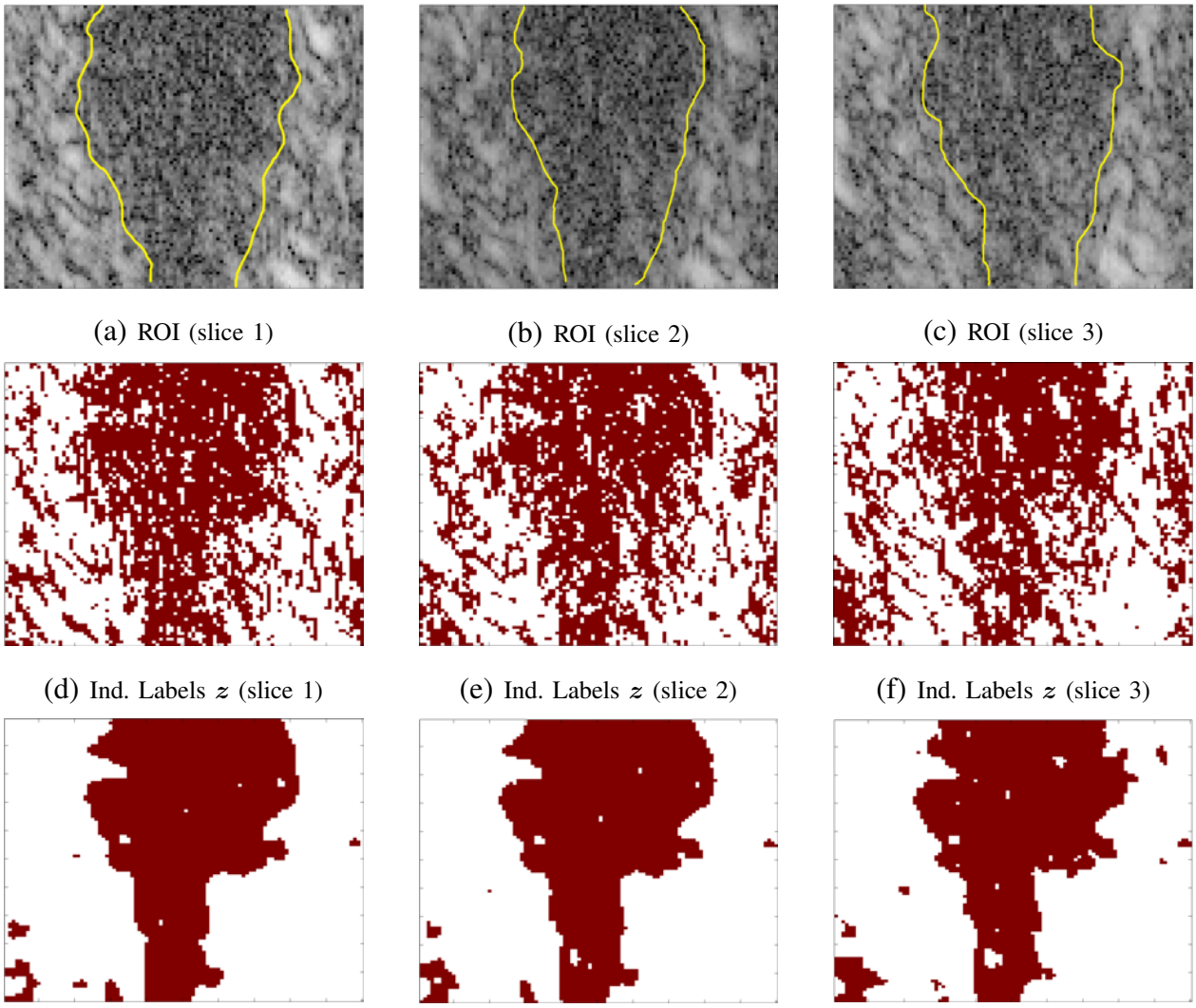
(d) Independent Labels $\boldsymbol{z}$

Fig. 8. Log-compressed US images of skin lesion and the corresponding estimated labels (healthy = white, lesion = red) [34])

.

In a second experiment the algorithm was applied in three dimensions using a tridimensional random field. Three slices of the 3D B-mode image associated with the ROI are shown in Figs. 9(a), 9(b) and 9(c). Figs. 9(d), 9(e) and 9(f) show the results obtained when labels are considered a priori independent, as in [34]. The labels estimated with the proposed 3D method are displayed in Figs. 9(g), 9(h) and 9(i) where healthy voxels are represented in white and lesion voxels in red. The size of the 3D images is $100 \times 100 \times 3$ voxels and computing class label estimates using

$1,000$ iterations of the proposed algorithm required 43.5 seconds (see Section V-C7 for more details about the computational complexity). We observe that most of the MAP labels are in very good agreement with the expert annotations. The improvement obtained when considering correlations in the 3rd dimension can be assessed by comparing Figs. 8(c) and 9(h), which have been computed from the same data slice. We observe that using a 3D MRF reduces significantly the number of misclassifications and improves the agreement with the expert annotations.



(a) ROI (slice 1) (b) ROI (slice 2) (c) ROI (slice 3)

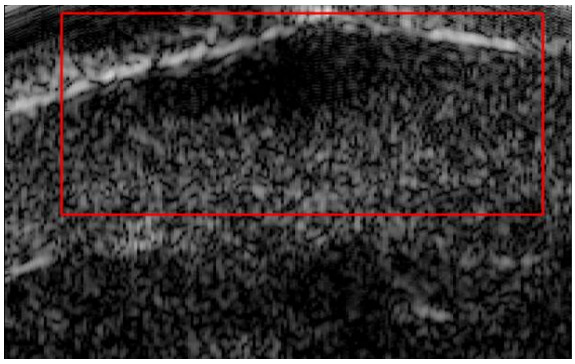
(d) Ind. Labels z (slice 1) (e) Ind. Labels z (slice 2) (f) Ind. Labels z (slice 3)

(g) MRF Labels z (slice 1) (h) MRF Labels z (slice 2) (i) MRF Labels z (slice 3)

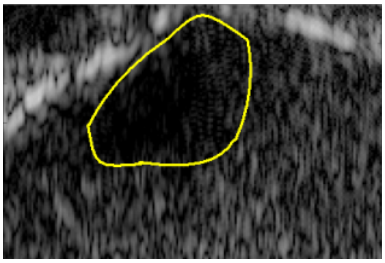
Fig. 9. Log-compressed US images of skin lesion and the corresponding estimated labels (*healthy* = white, *lesion* = red). Figs. (d)-(f) show the results obtained by considering that voxel labels are independent, as in [34]. Figs. (g)-(i) show the results obtained with the proposed 3D Markov random field (MRF) method.

3) Comparison with a state of the art method: The proposed algorithm has been compared with the state of the art method proposed in [23]. This method considers implicitly that the image is a mixture of two Rayleigh components and separates them using an LS algorithm. Comparison has been performed with 2D and 3D random fields. The following experiments were conducted with granularity coefficient $\beta = 1$ and number of classes $K = 4$ since there are 3 types of healthy tissue within the ROI in addition to the lesion. The results have been computed from a single Markov chain of $1,000$ iterations whose first 900 iterations (burn-in period) have been removed.

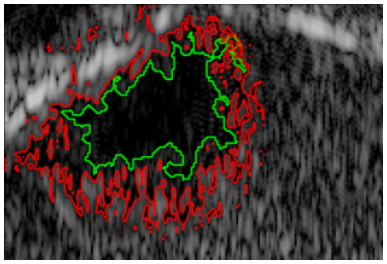
Fig. 10(a) shows a skin lesion contained in the ROI outlined by the red rectangle. This region is displayed with coarse expert annotations in Fig. 10(b). The proposed 2D Bayesian algorithm was used to label each voxel of the ROI as *healthy* or *lesion* tissue. Then, from the vector of voxels that were labeled as *lesion* we extracted the contour of the largest connected region. The results displayed in Fig. 10(c) show the regular shape of the contour obtained by our method, whereas the LS method with strong regularization yields a more irregular contour.



(a) Dermis view with skin lesion (ROI $= 160 \times 175 \times 16$).



(b) ROI (slice 7)



(c) 2D Segmentation contour

Fig. 10. Log-compressed US images of skin melanoma tumor and the corresponding estimated segmentation contours (*proposed* = green, [23] = red). Figure (c) proposed 2D algorithm and level set

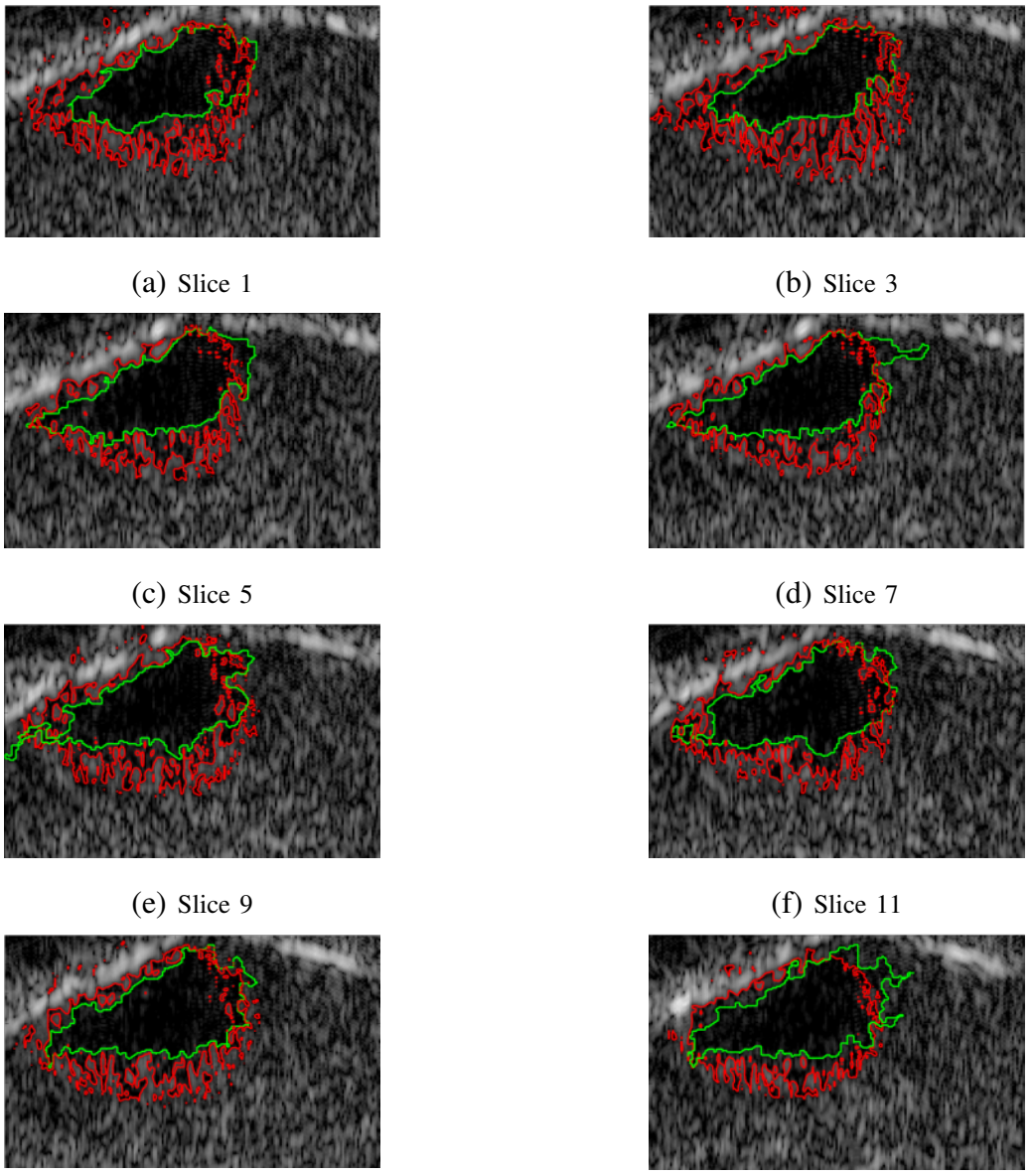
The proposed algorithm was also applied to a 3D B-mode image using a tridimensional random field. The results for eight slices of the image associated with the ROI depicted in Fig. 10(a) are shown in Figs. 11(a) - 11(h). The same color code is used for the contours as in the 2D experiment. The regular shape of the contour obtained by the proposed method is more visible and the recovered lesion fits better the area depicted by the expert. Finally, Fig. 12 shows two viewpoints of a 3D reconstruction of the lesion's surface. We observe that the tumor has a semi-ellipsoidal shape which is cut at the upper left by the epidermis-dermis junction. The tumor grows from this junction towards the deeper dermis, which is at the lower right.

Finally, it should be noted that in the in-vivo experiments the proposed algorithm has been applied to regions of interest, as opposed to entire 3D images. This has been motivated by the fact that dermatological ultrasound imaging is used to examine specific regions that have been previously identified by the dermatologist. The method presented in this work should be understood in that clinical context and is not intended to be used in unsupervised applications.

4) Segmentation results for different values of β: To assess the influence of the granularity coefficient, this section presents segmentation results obtained by repeating the previous experiment using different values of β. As explained previously, this hyper-parameter tunes the amount of correlation that the Potts Markov field introduces between the class labels. A small value of β defines a weak prior distribution that is very sensitive to noise, contrary to a large value of β that leads to a strong prior that promotes few and large homogeneous regions.

Fig. 13(a) shows the 8th slice of the B-mode 3D ROI previously displayed in Figs. 10 and 11. For visual interpretation this image is displayed in logarithmic scale. Figs. 13(b)-(f) show the 8th slice of the 3D MAP class labels obtained with the proposed algorithm for different values of β. These results were computed using $K = 4$ and are displayed using the following color code: *lesion* = black, *epidermis* = white, *pap. dermis* = dark gray and *ret. dermis* = light gray. These estimates have been computed from a single Markov chain of $1,000$ iterations whose first 900 iterations (burn-in period) have been removed.

We observe that the best results are obtained for $\beta = 1$ and $\beta = 1.25$. The results obtained by fixing the granularity coefficient to a small value ($\beta < 1$) are corrupted by ultrasound speckle noise and fail to capture the different skin layers. On the other hand, fixing β to a too high value (i.e., $\beta > 1.5$) enforces too much spatial correlation and yields a segmentation with artificially straight boundaries.



(a) Slice 1 (b) Slice 3

(c) Slice 5 (d) Slice 7

(e) Slice 9 (f) Slice 11

(g) Slice 13 (h) Slice 15

Fig. 11. 3D segmentation of an 8-slice image.

5) Segmentation results for different numbers of classes K: As explained previously, this work assumed that the number of classes K is known a-priori. For skin tissues the number of classes depends on the number of layers contained in that region (i.e., epidermis, papillary (upper) dermis, reticular (lower) dermis, hypodermis) in addition to the lesion. The number of classes will typically vary from $K = 3$ for very small lesions (contained in the upper dermis) to

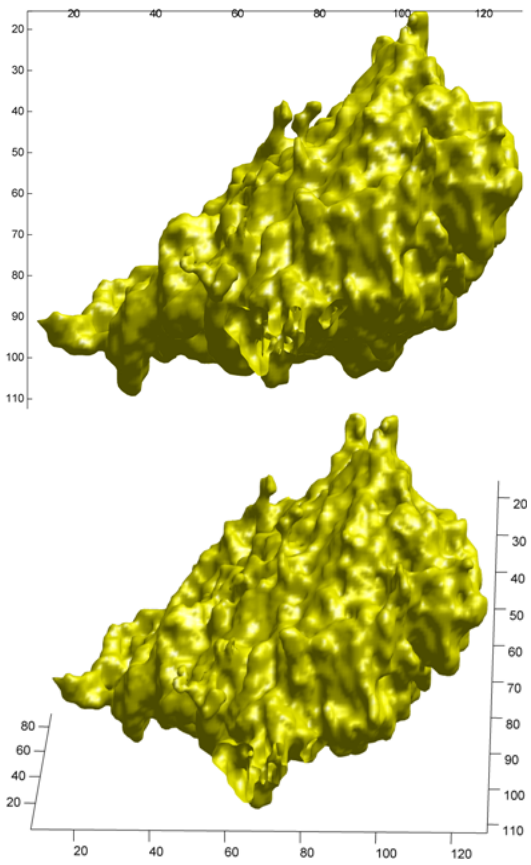
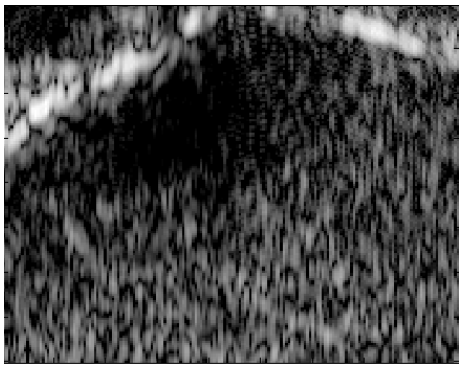


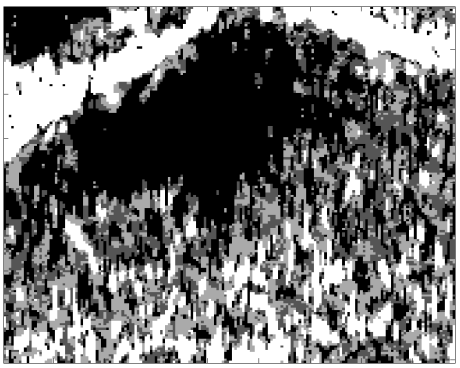
Fig. 12. 3D reconstruction of the melanoma tumor.

$K = 5$ for lesions that have invaded the lower dermis. In any case the number of classes should be at least equal to the number of tissues that must be identified.

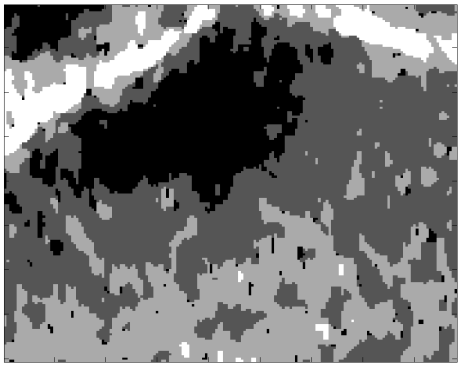
Note that increasing the number of classes beyond the number of tissues will outline differences within a same biological tissue. To illustrate this point the proposed method has been applied to one of our 3D ultrasound images using one additional class. Fig. 14(b) shows one slice of the MAP labels estimated using one class per tissue (one class per skin layer plus one class for the lesion), i.e., $K = 4$. Fig. 14(c) shows results obtained when considering an additional class, i.e., $K = 5$. We observe that introducing an additional class has not modified significantly the estimation of the lesion boundaries. The proposed method has assigned an additional class to the core of the lesion, which may correspond to a necrotic tissue. Moreover, using too many additional classes will result in empty or redundant classes. Fig. 14(d) shows one slice of the MAP labels obtained with two additional classes ($K = 6$). We observe that this result is very similar to the one obtained for a single additional class ($K = 5$). In this case the proposed method



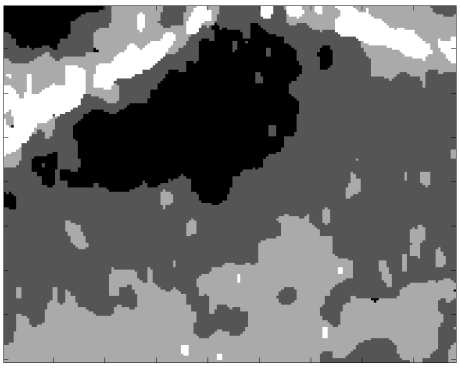
(a) Dermis view with skin lesion (slice 8)



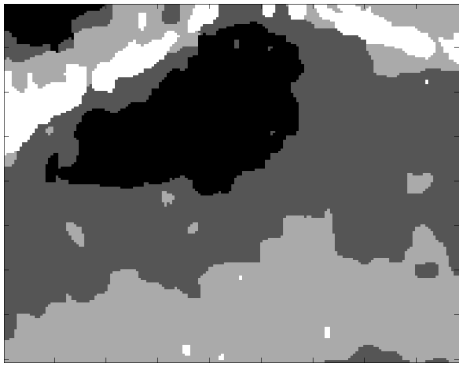
(b) MAP Class Labels ($\beta = 0.5$)



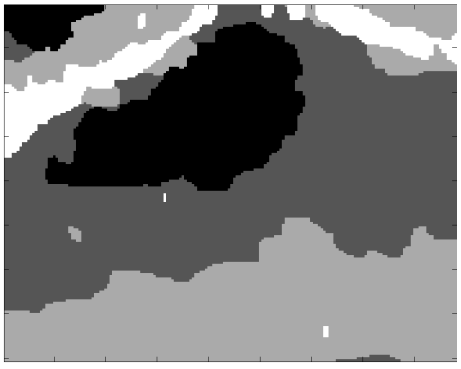
(c) MAP Class Labels ($\beta = 0.75$)



(d) MAP Class Labels ($\beta = 1.0$)



(e) MAP Class Labels ($\beta = 1.25$)



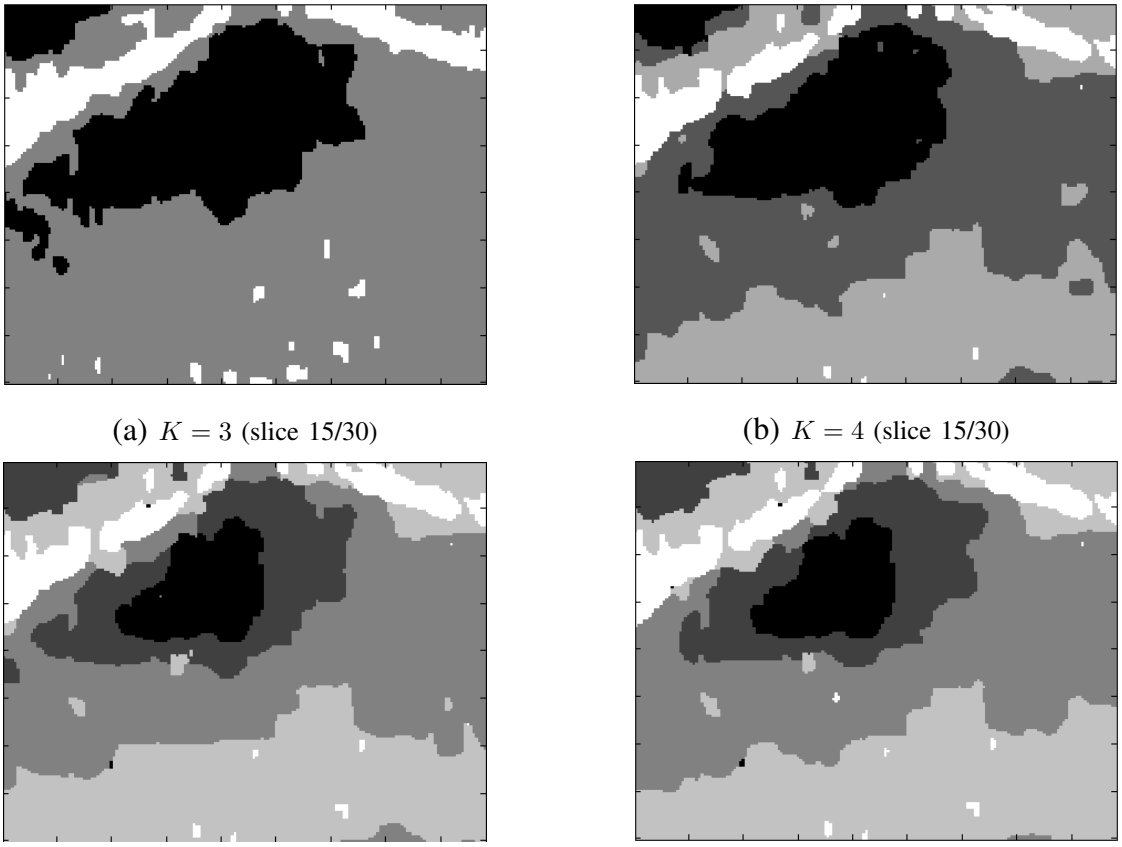
(f) MAP Class Labels ($\beta = 1.5$)

Fig. 13. Log-compressed US images of skin lesion and the corresponding estimated class labels (*lesion* = black, *epidermis* = white, *pap. dermis* = dark gray, *ret. dermis* = light gray)

has assigned one additional class to the core of the lesion and left the other supplementary class unassigned, indicating that $K = 5$ is a more suitable number of classes.

On the other hand, underestimating the number of classes can degrade the performance of the

proposed method significantly. Fig. 14(a) shows one slice of the segmentation results obtained when there are not enough classes to represent all tissues, i.e., $K = 3$. In this case the papillary dermis and the reticular dermis have been merged into a single class. As a result the estimation of the boundary between the lesion and the papillary dermis is less accurate.



(a) $K = 3$ (slice 15/30)

(b) $K = 4$ (slice 15/30)

(c) $K = 5$ (slice 15/30)

(d) $K = 6$ (slice 15/30)

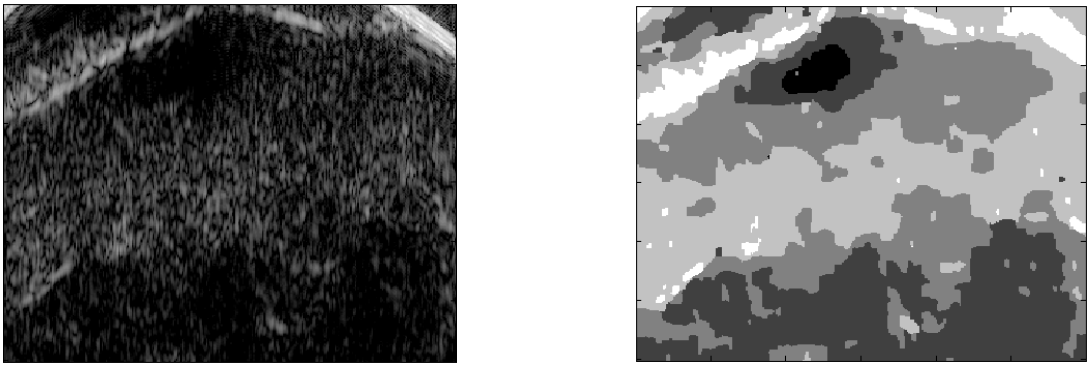
Fig. 14. Tissue labeling results (central slice) of a 3D ultrasound image containing a lesion. (a) Using too few classes results in a common class for papillary and reticular dermis. (b) Correctly using one class per tissue. (c) The introduction of a supplementary class reveals the core of the lesion. (d) The introduction of two supplementary classes results in an empty class.

6) Segmentation of entire 3D B-mode images: In this work the proposed algorithm has been applied to regions of interest, as opposed to entire 3D images. This is motivated by the fact that dermatological ultrasound is used to reexamine carefully regions that have been previously identified using a faster screening modality, typically a dermatoscope (magnifying glass). The method presented in this work should be understood in that clinical context and is not intended

for batch processing entire 3D images nor for fully unsupervised applications.

However, for completeness the proposed method has also been applied to an entire B-mode ultrasound image. It should be noted that identifying the hypodermis (the lower part of the image) is difficult because the ultrasound system is calibrated to target the upper dermis, which is where lesions are commonly found. The hypodermis is significantly more affected by attenuation, resolution degradation and noise.

Figure 15 shows one slice of the ultrasound image and the corresponding 3D MAP class labels estimated with the proposed algorithm. For visual interpretation Fig. 15(a) is displayed using logarithmic scale. The results shown in Fig. 15(b) have been computed using $K = 5$ classes because there are 4 types of healthy tissue in addition to the lesion. We observe that the lesion and its core have been accurately detected. Also, the boundaries between the skin layers have also been correctly identified. The hypodermis has been mostly associated to the same class as the lesion, probably because both tissues are very hypoechogenic. Increasing the number of classes did not improve this result.



(a) Log-compressed Ultrasound image (8th slice) (b) Labels (8th slice)

Fig. 15. Log-compressed US images of skin lesion and the corresponding estimated class labels computed on the entire image

7) Computational Complexity: Table II provides averaged execution times for 500 iterations of the proposed algorithm for several image sizes in 2D and 3D and several numbers of classes. The time required to reach convergence can be calculated by multiplying these values by $\frac{9}{5}$, which corresponds to a burn-in period of 900 iterations. These tests have been computed on a workstation equipped with an Intel Core 2 Duo @2.1 GHz processor, 3MB L2 and 3GB of RAM

memory. The main loop of the Gibbs sampler has been implemented on MATLAB R2010b (The MathWorks Inc., Natick, MA, 2010). However, C-MEX functions have been used to compute the likelihood and to draw samples of $\boldsymbol{z}$ from (15).

TABLE II

COMPUTING TIMES (IN SECONDS) OF 500 ITERATIONS FOR DIFFERENT IMAGE SIZES AND NUMBER OF CLASSES.

	K = 2	K = 4	K = 8	K = 16
64×64	4.9	10.2	18.5	35.0
128×128	8.1	14.6	27.0	51.7
256×256	19.7	36.1	63.5	123.0
512×512	73.7	126.0	223.2	427.0
$64 \times 64 \times 16$	20.8	36.6	68.3	129.5
$128 \times 128 \times 16$	75.1	141.5	254.0	524.4
$256 \times 256 \times 16$	317.7	578.1	1060.5	2300.5
$512 \times 512 \times 16$	1175.7	2225.0	4316.8	9600.0

Finally, table III provides the average computing times for the LS method [23] for different image sizes. These results have been computed using the time and space sampling steps indicated in [23] ($\Delta t = 0.1$, $\Delta x = 1$, $\Delta y = 1$ and $\Delta z = 1$). Average estimates were obtained by repeating each experiment 10 times. Note that a comparison between these computing times and those of the proposed method has to be made with some precautions since the LS method has been implemented in [23] using MATLAB whereas the proposed method uses C-MEX functions. .

TABLE III

AVERAGE COMPUTING TIMES FOR THE LEVEL SET METHOD [23] FOR DIFFERENT IMAGE SIZES.

	250 iterations	time to convergence
64×64	0.77 sec.	0.77 sec. (250 iterations)
128×128	2.38 sec.	9.52 sec. (1000 iterations)
256×256	12.23 sec.	73.35 sec. (1500 iterations)
512×512	22.95 sec.	183.6 sec. (2000 iterations)

VI. Conclusion

A spatially coherent finite mixture of α-Rayleigh distributions was proposed to represent the statistics of envelope ultrasound images backscattered from multiple tissues. Spatial correlation was introduced into the model by a Markov random field that promotes dependance between neighbor pixels. Based on the proposed model, a Bayesian segmentation method was derived. Bidimensional and tridimensional implementations of this segmentation method were presented using a Markov chain Monte Carlo algorithm that jointly estimates the unknown parameters of the mixture model and classifies voxels into different tissues. The method was successfully applied to several high frequency 3D ultrasound images. Experimental results showed that the proposed technique outperforms a state of the art method in the segmentation of in-vivo lesions. A tridimensional reconstruction of a melanoma tumor suggested that the resulting segmentations can be used to assess lesion penetration in dermatologic oncology. Future work includes the characterization of the performance of the segmentation algorithm and the study of estimation algorithms for the granularity coefficient defining the Markov random field prior. A comparison with an ML estimator followed by median filtering is also an area of interest for potential future work.

Acknowledgment

This work was developed as part of the CAMM4D project, supported by the French FUI and the Midi Pyrenees Regional Council. We would like to acknowledge the support of the Hospital of Toulouse, Pierre Fabre Laboratories and Magellium for the ultrasound image acquisition. We would particularly like to thank Dr. Nicolas Meyer, Dr. Siham Lourari, and Mr. Jérôme Georges.

References

[1] J. Noble and D. Boukerroui, "Ultrasound image segmentation: a survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[2] A. Belaid, D. Boukerroui, Y. Maingourd, and J.-F. Lerallut, "Phase-based level set segmentation of ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 1, pp. 138–147, jan. 2011.

[3] R. Schneider, D. Perrin, N. Vasilyev, G. Marx, P. del Nido, and R. Howe, "Mitral annulus segmentation from 3d ultrasound using graph cuts," *IEEE Trans. Med. Imag.*, vol. 29, no. 9, pp. 1676–1687, sept. 2010.

[4] K. Somkantha, N. Theera-Umpon, and S. Auephanwiriyakul, "Boundary detection in medical images using edge following algorithm based on intensity gradient and texture gradient features," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 567–573, march 2011.

[5] P. Yan, S. Xu, B. Turkbey, and J. Kruecker, "Adaptively learning local shape statistics for prostate segmentation in ultrasound," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 633–641, march 2011.

[6] G. Unal, S. Bucher, S. Carlier, G. Slabaugh, T. Fang, and K. Tanaka, "Shape-driven segmentation of the arterial wall in intravascular ultrasound images," *IEEE Trans. Inf. Technol. Biomed.*, vol. 12, no. 3, pp. 335–347, may 2008.

[7] G. Carneiro, B. Georgescu, S. Good, and D. Comaniciu, "Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree," *IEEE Trans. Med. Imag.*, vol. 27, no. 9, pp. 1342–1355, sept. 2008.

[8] P. Paramanathan and R. Uthayakumar, "Tumor growth in the fractal space-time with temporal density," in *Control, Computation and Information Systems*, ser. Communications in Computer and Information Science. Springer Berlin Heidelberg, 2011, vol. 140, pp. 166–173.

[9] K. Horsch, M. L. Giger, L. A. Venta, and C. J. Vyborny, "Automatic segmentation of breast lesions on ultrasound," *Medical Physics*, vol. 28, no. 8, pp. 1652–1659, 2001.

[10] I.-S. Jung, D. Thapa, and G.-N. Wang, "Automatic segmentation and diagnosis of breast lesions using morphology method based on ultrasound," in *Fuzzy Systems and Knowledge Discovery*, ser. Lecture Notes in Computer Science, L. Wang and Y. Jin, Eds. Springer Berlin / Heidelberg, 2005, vol. 3614, pp. 491–491.

[11] A. Madabhushi and D. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 155–169, feb. 2003.

[12] Y.-L. Huang and D.-R. Chen, "Watershed segmentation for breast tumor in 2-d sonography," *Ultrasound in Medicine and Biology*, vol. 30, no. 5, pp. 625–632, 2004.

[13] X. Guofang, M. Brady, J. Noble, and Z. Yongyue, "Segmentation of ultrasound b-mode images with intensity inhomogeneity correction," *IEEE Trans. Med. Imag.*, vol. 21, no. 1, pp. 48–57, jan. 2002.

[14] D. Boukerroui, A. Baskurt, J. Noble, and O. Basset, "Segmentation of ultrasound images–multiresolution 2d and 3d algorithm based on global and local statistics," *Pattern Recog. Letters*, vol. 24, no. 4-5, pp. 779–790, 2003.

[15] J. Dias and J. Leitao, "Wall position and thickness estimation from sequences of echocardiographic images," *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 25–38, feb 1996.

[16] R. Wagner, S. Smith, J. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound b-scans," *IEEE Trans. Sonics and Ultrasonics*, vol. 30, no. 3, pp. 156–163, May 1983.

[17] E. Brusseau, C. de Korte, F. Mastik, J. Schaar, and A. van der Steen, "Fully automatic luminal contour segmentation in intracoronary ultrasound imaging-a statistical approach," *IEEE Trans. Med. Imag.*, vol. 23, no. 5, pp. 554–566, may 2004.

[18] C. Chesnaud, P. Refregier, and V. Boulet, "Statistical region snake-based segmentation adapted to different physical noise models," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1145–1157, nov 1999.

[19] M. H. R. Cardinal, J. Meunier, G. Soulez, R. L. Maurice, E. Therasse, and G. Cloutier, "Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions," *IEEE Trans. Med. Imag.*, vol. 25, no. 5, pp. 590–601, May 2006.

[20] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Physics*, vol. 79, no. 1, pp. 12–49, 1988.

[21] L. Saroul, O. Bernard, D. Vray, and D. Friboulet, "Prostate segmentation in echographic images: A variational approach using deformable super-ellipse and Rayleigh distribution," in *Proc. IEEE Int. Symp. Biomed. Imag (ISBI)*, may 2008, pp. 129–132.

[22] F. Destrempes, J. M. J., M. F. Giroux, G. Soulez, and G. Cloutier, "Segmentation in ultrasonic B-mode images of carotid

arteries using mixture of Nakagami distributions and stochastic optimization," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 215–229, Feb. 2009.

[23] A. Sarti, C. Corsi, E. Mazzini, and C. Lamberti, "Maximum likelihood segmentation of ultrasound images with rayleigh distribution," *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 52, no. 6, pp. 947–960, june 2005.

[24] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, feb 2001.

[25] F. Lecellier, J. Fadili, S. Jehan-Besson, G. Aubert, M. Revenu, and E. Saloux, "Region-based active contours with exponential family observations," *J. Math. Imaging Vis.*, vol. 36, pp. 28–45, January 2010.

[26] M. A. Pereyra and H. Batatia, "Modeling ultrasound echoes in skin tissues using symmetric α-stable processes," *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 59, no. 1, pp. 60–72, Jan. 2012.

[27] F. Y. Wu, "The Potts model," *Rev. Mod. Phys.*, vol. 54, no. 1, pp. 235–268, Jan. 1982.

[28] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4239–4247, Nov. 2011.

[29] D. Van de Sompel and M. Brady, "Simultaneous reconstruction and segmentation algorithm for positron emission tomography and transmission tomography," in *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, may 2008, pp. 1035 –1038.

[30] B. Scherrer, F. Forbes, C. Garbay, and M. Dojat, "Distributed local MRF models for tissue and structure brain segmentation," *IEEE Trans. Med. Imag.*, vol. 28, no. 8, pp. 1278 –1295, aug. 2009.

[31] L. Risser, J. Idier, P. Ciuciu, and T. Vincent, "Fast bilinear extrapolation of 3D Ising field partition function. Application to fMRI image analysis," in *Proc. ICIP*, Cairo, Egypte, Nov. 2009, pp. 833–836.

[32] J. Marroquin, E. Santana, and S. Botello, "Hidden markov measure field models for image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 25, no. 11, pp. 1380 – 1387, nov. 2003.

[33] M. Woolrich, T. Behrens, C. Beckmann, and S. Smith, "Mixture models with adaptive spatial regularization for segmentation with an application to fMRI data," *IEEE Trans. Med. Imag.*, vol. 24, no. 1, pp. 1 –11, jan. 2005.

[34] M. A. Pereyra, N. Dobigeon, H. Batatia, and J.-Y. Tourneret, "Labeling skin tissues in ultrasound images using a generalized Rayleigh mixture model," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, Prague, Czech Republic, May 2011, pp. 729–732.

[35] ——, "Segmentation of ultrasound images using a spatially coherent generalized Rayleigh mixture model," in *European Signal Processing Conf. (EUSIPCO)*, Barcelona, Spain, Sep 2011, pp. 664–668.

[36] P. Morse and K. Ingard, *Theoretical Acoustics*. Princeton (NJ): Princeton University Press, 1987.

[37] D. Salas-Gonzalez, E. E. Kuruoglu, and D. P. Ruiz, "Finite mixture of α-stable distributions," *Digit. Signal Process.*, vol. 19, pp. 250–264, March 2009.

[38] E. E. Kuruoglu and J. Zerubia, "Modeling SAR images with a generalization of the Rayleigh distribution," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 527 – 533, April 2004.

[39] A. Achim, E. E. Kuruoglu, and J. Zerubia, "SAR image filtering based on the heavy-tailed Rayleigh model," *IEEE Trans. on Image Processing*, vol. 15, no. 9, pp. 2686–2693, September 2006.

[40] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[41] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Stat. Soc. Ser. B*, vol. 36, no. 2, pp. 192–236, 1974.

[42] R. Kindermann and J. L. Snell, *Markov random fields and their applications*. Providence: RI: Amer. Math. Soc., 1980.

[43] Z. Zhou, R. Leahy, and J. Qi, "Approximate maximum likelihood hyperparameter estimation for Gibbs prior," *IEEE Trans. Image Process.*, vol. 6, no. 6, pp. 844–861, June 1997.

[44] X. Descombes, R. Morris, J. Zerubia, and M. Berthod, "Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood," *IEEE Trans. Image Process.*, vol. 8, no. 7, pp. 945–963, June 1999.

[45] J. Moller, A. N. Pettitt, R. Reeves, and K. K. Berthelsen, "An efficient Markov chain Monte Carlo method for distributions with intractable normalising constants," *Biometrika*, vol. 93, no. 2, pp. 451–458, June 2006.

[46] C. McGrory, D. Titterington, R. Reeves, and A. Pettitt, "Variational Bayes for estimating the parameters of a hidden Potts model," *Statistics and Computing*, vol. 19, pp. 329–340, 2009.

[47] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer-Verlag, 1999.

[48] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Stat. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.

[49] J. Diebolt and E. H. S. Ip., "Stochastic EM: method and application," in *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds. London: Chapman & Hall, 1996.

[50] N. Dobigeon and J.-Y. Tourneret, "Bayesian orthogonal component analysis for sparse representation," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2675–2685, May 2010.

[51] N. Dobigeon, A. Hero, and J.-Y. Tourneret, "Hierarchical Bayesian sparse image reconstruction with application to MRFM," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 2059–2070, Sept. 2009.

[52] T. Vincent, L. Risser, and P. Ciuciu, "Spatially adaptive mixture modeling for analysis of fMRI time series," *IEEE Trans. Med. Imag.*, vol. 29, no. 4, pp. 1059–1074, April 2010.

[53] K. Kayabol, E. Kuruoglu, and B. Sankur, "Bayesian separation of images modeled with MRFs using MCMC," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 982–994, May 2009.

[54] M. Mignotte, "Image denoising by averaging of piecewise constant simulations of image partitions," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 523–533, Feb. 2007.

[55] J. Gonzalez, Y. Low, A. Gretton, and C. Guestrin, "Parallel Gibbs sampling: From colored fields to thin junction trees," in *Artificial Intelligence and Statistics (AISTATS)*, Ft. Lauderdale, FL, May 2011.

[56] G. O. Roberts, "Markov chain concepts related to sampling algorithms," in *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds. London: Chapman & Hall, 1996, pp. 259–273.

[57] Z. Sun and C. Han, "Heavy-tailed Rayleigh distribution: A new tool for the modeling of SAR amplitude images," in *Proc. (IGARSS 08). IEEE Int. Geosc. and Remote Sensing Symp.*, vol. 4, July 2008, pp. 1253–1256.

[58] C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distribution and Applications*. New York (NJ): Wiley, 1995.

[59] A. Gelman and D. Rubin, "Inference from iterative simulation using multiple sequences," *Stat. Sciences*, vol. 7, no. 4, pp. 457–511, 1992.

[60] C. P. Robert and S. Richardson, "Markov Chain Monte Carlo methods," in *Discretization and MCMC Convergence Assessment*, C. P. Robert, Ed. New York: Springer Verlag, 1998, pp. 1–25.

[61] A. Gelman, J. B. Carlin, H. P. Robert, and D. B. Rubin, *Bayesian Data Analysis*. London: Chapman & Hall, 1995.

[62] T. Hergum, S. Langeland, E. Remme, and H. Torp, "Fast ultrasound imaging simulation in k-space," *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 56, no. 6, pp. 1159 –1167, june 2009.

[63] T. Szabo, *Diagnostic Ultrasound Imaging*. Boston,USA: Academic Press, 2004.