LABELING SKIN TISSUES IN ULTRASOUND IMAGES USING A GENERALIZED RAYLEIGH MIXTURE MODEL

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

University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse Cedex 7, France

{marcelo.pereyra, nicolas.dobigeon, hadj.batatia, jean-yves.tourneret}@enseeiht.fr

ABSTRACT

This paper addresses the problem of estimating the statistical distribution of multiple-tissue non-stationary ultrasound images of skin. The distribution of multiple-tissue images is modeled as a finite mixture of Heavy-Tailed Rayleigh distributions. An original Bayesian algorithm combined with a Markov chain Monte Carlo method is then derived to jointly estimate the mixture parameters and a label vector associating each voxel to a tissue. Precisely, a hybrid Metropolis-within-Gibbs sampler is proposed to draw samples that are asymptotically distributed according to the posterior distribution of the Bayesian model. These samples are then used to compute the Bayesian estimators of the model parameters. Simulation results are conducted on synthetic data to illustrate the performance of the proposed estimation strategy. The method is then successfully applied to the detection of an in-vivo skin lesion in a high frequency 3D ultrasound image.

Index Terms— Heavy-tailed Rayleigh distribution, mixture model, Bayesian estimation, Gibbs sampler.

1. INTRODUCTION

It has long been acknowledged that ultrasound speckle is an information resource that can be exploited for tissue characterization [11][8]. In this context, biological tissue characterization is achieved by analyzing the statistical distribution of the backscattered ultrasound signal. This calls for statistical analysis models that are both coherent with the observed statistics and with the underlying signal formation physics [11][9].

Statistical studies of ultrasound scattering in biological tissues are generally conducted within the widely accepted point scattering framework [7]. It has been recently established that under this representation, the radio-frequency (RF) signal backscattered from skin tissues converges in distribution towards a non-Gaussian symmetric α-stable law ($S\alpha S$)[10]. Similarly, the envelope signal (or B-mode) converges to a generalized heavy-tailed Rayleigh distribution denoted as α-Rayleigh distribution. This theoretical result is in accordance with other studies that have empirically shown that signals backscattered from skin and breast tissues are well described by stable distributions [9][6].

This paper extends the single-tissue stationary α-stable model to multiple-tissue piecewise stationary signals. Accordingly, instead of a single α-Rayleigh distribution, we propose to represent the distribution of the envelope signal as a mixture of α-Rayleigh distributions. In addition, we derive a Bayesian algorithm combined with a Markov chain Monte Carlo (MCMC) method for inferring the mixture parameters.

2. PROBLEM STATEMENT

Let $r_n \in \mathbb{R}^+$ denote an observation, or voxel, in an envelope ultrasound image $\boldsymbol{r} \in \mathbb{R}^N$. In addition, let r_n be defined by means of the widely accepted point scattering model [7]

$$r_n = \left| \sum_{i=0}^{M_n} a_i \exp\left[\jmath\left(\omega_0 t_n + \theta_i\right)\right] \right| \tag{1}$$

where t_n is the propagation delay, ω_0 is the transducer central frequency, M_n is the number of scatterers in the nth resolution cell, and the random variables a_i and θ_i are the amplitude and phase of the wave backscattered from the ith scatterer within that resolution cell. Recent works on scattering in biological tissues have established that r_n, as defined above, converges in distribution towards an α-Rayleigh distribution [10]

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

where $\underset{M_n \to \infty}{\overset{d}{\to}}$ denotes convergence in distribution when M_n approaches ∞, and the parameters $\alpha_n \in (0, 2]$ and $\gamma_n \in \mathbb{R}^+$ are the characteristic index and spread associated to the nth resolution cell.

This paper considers the case where the ultrasound image $\boldsymbol{r} = (r_1, \ldots, r_N)^T$ is made up by multiple biological tissues, 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 $\{C_1, ..., C_K\}$ such that

$$\forall r_n \in \mathcal{C}_k, \quad r_n \underset{M_n \to \infty}{\overset{d}{\to}} \alpha\mathcal{R}(\alpha_k, \gamma_k) \tag{3}$$

where α_k and γ_k are the parameters associated to the class $\mathcal{C}_k$. As a consequence, it is possible to approximate the distribution of an observation within $\boldsymbol{r}$ by means of the following mixture of α-Rayleigh distributions

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

where K is the number of classes and w_k represents the relative weight (or proportion) of the kth class. It should be noted here that the proposed α-Rayleigh mixture model generalizes the Rayleigh mixture model, which has been extensively applied to multiple-tissue ultrasound image modeling [15] [1]. At last, we acknowledge the existence of a previous generalization of the Rayleigh mixture, the Nakagami mixture [2], which would be however unsuitable for skin tissues that exhibit strong non-Nakagami statistics [11].

3. BAYESIAN MODEL

This section addresses the problem of estimating the parameters of the α-Rayleigh mixture model introduced in (4). The unknown parameter vector for this problem can be written $\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \boldsymbol{\gamma}^T)^T$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_K)^T$, $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)^T$. In addition, we explicitly introduce the hidden label vector $\boldsymbol{z} = \{z_1, \ldots, z_N\}$ associated to the observation vector $\boldsymbol{r} = \{r_1, \ldots, r_N\}$ such that $z_n = k$ if $r_n \in C_k$. Note that the prior probabilities of having $z_n = k$, $k = 1, \ldots, K$, are $\mathrm{P}[z_n = k] = \omega_k$, where ω_k ($k = 1, \ldots, K$) are the mixture weights in (4). These labels will allow one to characterize each image observation and, possibly, to discriminate the different kinds of tissue. The number of classes K is assumed to be known in this study. This assumption might be relaxed by studying a reversible jump MCMC algorithm [5, 12]. This section defines a Bayesian model associated to the unknown parameter vector $\boldsymbol{\theta}$ and to the unknown label vector $\boldsymbol{z}$. This model requires to define the likelihood of the observation vector $\boldsymbol{r}$ and the priors for the unknown parameters.

3.1. 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

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

where

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

and where J_0 is the 0th order Bessel function of the first kind.

3.2. Parameter priors

3.2.1. Labels

Since there is no prior information regarding the different classes $C_1, \ldots, C_K$, it is natural to assume that all the values over the set $\{1, \ldots, K\}$ are a priori equiprobable for each label z_n

$$\mathrm{P}[z_n = k] = \frac{1}{K}, \quad \forall k = 1, \ldots, K. \tag{7}$$

Assuming the labels $z_1, \ldots, z_N$ are a priori independent, the joint prior probability for the label vector $\boldsymbol{z}$ is

$$\mathrm{P}[\boldsymbol{z}] = \prod_{n=1}^{N} \mathrm{P}[z_n]. \tag{8}$$

3.2.2. α-Rayleigh parameters

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

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

since the interval $(0, 2]$ covers all possible values of this parameter.

The prior for the α-Rayleigh spread γ_k is an inverse gamma distribution with hyperparameters a_0 and b_0

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

where the hyperparameters are fixed to $a_0 = 1$ and $b_0 = 1$ yielding a vague prior.

Assuming that each of these α-Rayleigh parameters are a priori independent, the prior distribution for the parameter vector $\boldsymbol{\theta}$ is

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

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

Assuming the unknown parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{z}$ are a priori independent and using Bayes theorem, the posterior distribution of the parameter vector $(\boldsymbol{\theta}, \boldsymbol{z})$ can be expressed as follows

$$p(\boldsymbol{\theta}, \boldsymbol{z}|\boldsymbol{r}) \propto p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z})p(\boldsymbol{\theta})p(\boldsymbol{z}) \tag{12}$$

where the likelihood $p(\boldsymbol{r}|\boldsymbol{\theta}, \boldsymbol{z})$ and the prior distribution $p(\boldsymbol{\theta})$ have been defined in (5) and (11), respectively.

Unfortunately the posterior distribution (12) is too complex to derive the minimum mean square error (MMSE) or maximum a posteriori (MAP) estimators of the unknown parameters $\boldsymbol{\theta}$ and $\boldsymbol{z}$. Note that the expectation-maximization (EM) algorithm, which is frequently used for finite mixture problems [15, 2] is not appropriate in the context of α-Rayleigh distributions since the expectation of (12) is not defined if there is at least one characteristic index $\alpha_k < 1$. An interesting alternative is to use an MCMC method to generate samples that are asymptotically distributed according to the target distribution (12) [13]. Then the generated samples can be used to approximate the Bayesian estimators, as in [3].

4. HYBRID GIBBS SAMPLER

This section studies a hybrid Metropolis-within-Gibbs sampler that draws samples that are asymptotically distributed according to (12). The principle is to simulate according to the conditional distributions of the target distribution that are provided below. The interested reader is invited to consult [13] for more details about MCMC methods.

4.1. Conditional probability $\mathrm{P}[\boldsymbol{z}|\boldsymbol{\theta}, \boldsymbol{r}]$

The label vector $\boldsymbol{z}$ can be updated coordinate-by-coordinate using Gibbs move. More precisely, the conditional probabilities $\mathrm{P}[z_n|\boldsymbol{\theta}, \boldsymbol{r}]$ can be computed following the Bayes rule

$$\begin{aligned} \mathrm{P}[z_n = k|r_n, \alpha_k, \gamma_k] \propto \pi_{n,k} &\triangleq p_{\alpha\mathcal{R}}(r_n|z_n = k, \alpha_k, \gamma_k)\mathrm{P}[z_n = k] \\ &= \frac{1}{K} p_{\alpha\mathcal{R}}(r_n|z_n = k, \alpha_k, \gamma_k) \end{aligned} \tag{13}$$

where $\propto$ means "proportional to" and $p_{\alpha\mathcal{R}}(\cdot)$ has been defined in (6). Once all the quantities $\pi_{n,k}$, $k = 1, \ldots, K$, have been computed, they are normalized to obtain the posterior probabilities $\tilde{\pi}_{n,k} \triangleq \mathrm{P}[z_n = k|r_n, \alpha_k, \gamma_k]$, i.e.,

$$\tilde{\pi}_{n,k} = \frac{\pi_{n,k}}{\sum_{k=1}^{K} \pi_{n,k}}. \tag{14}$$

Finally, samples z_n are generating by drawing discrete variables from $\{1, \ldots, K\}$ with the respective probabilities $\{\pi_{n,1}, \ldots, \pi_{n,K}\}$.

4.2. Conditional distribution of $p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{r})$

The generation of samples according to $p(\boldsymbol{\alpha}|\boldsymbol{\gamma}, \boldsymbol{z}, \boldsymbol{r})$ can be achieved coordinate-by-coordinate using a Metropolis-Hastings (MH) step, leading to a Metropolis-within-Gibbs algorithm [13, p. 317]. In this work, the uniform prior distribution defined in (9) has been chosen as proposal distribution for this algorithm. Given that both the prior and the proposal distributions are uniform, the MH acceptance rate of the proposed move is equal to the likelihood ratio

$$\text{ratio} = \min\left\{1, \prod_{n:z_n=k}^{N} \frac{p_{\alpha\mathcal{R}}(r_n|\alpha_k^*, \gamma_k)}{p_{\alpha\mathcal{R}}(r_n|\alpha_k^{(t-1)}, \gamma_k)}\right\} \tag{15}$$

where α_k^* denotes the proposed value at iteration t and $\alpha_k^{(t-1)}$ is the previous state of the chain.

4.3. Conditional distribution of $p(\boldsymbol{\gamma}|\boldsymbol{\alpha}, \boldsymbol{z}, \boldsymbol{r})$

Similarly, generation of samples according to $p(\boldsymbol{\gamma}|\boldsymbol{\alpha}, \boldsymbol{z}, \boldsymbol{r})$ is also achieved thanks to an MH step for each coordinate γ_k, successively. In this case, the proposal distribution is a normal distribution centered on the previous value of the chain with variance σ^2

$$\gamma_k^* \sim \mathcal{N}(\gamma_k^{(t-1)}, \sigma^2). \tag{16}$$

The hyperparameter σ^2 is chosen to ensure an acceptance ratio close to $\frac{1}{2}$, as recommended in [14]. Note that the proposal (16) results from the so-called random walk MH algorithm [13, p. 245]. Furthermore, since the proposal distribution is symmetric, the acceptance ratio can be computed directly from the likelihood and the prior ratios

$$\text{ratio} = \min\left\{1, \prod_{n:z_n=k}^{N} \frac{p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k^*)p(\gamma_k^*|a_0, b_0)}{p_{\alpha\mathcal{R}}(r_n|\alpha_k, \gamma_k^{(t-1)})p(\gamma_k^{(t-1)}|a_0, b_0)}\right\} \tag{17}$$

where the prior distribution $p(\gamma_k|a_0, b_0)$ has been defined in (10).

5. SIMULATIONS RESULTS

This section presents simulated and experimental results conducted to assess the performance of the proposed α-Rayleigh mixture model and the associated Bayesian estimation algorithm.

5.1. Synthetic Data

The purpose of the first simulation is to illustrate the method's capacity to resolve a mixture with components that are close to each other. Accordingly, the proposed algorithm was applied to a synthetic 2-component α-Rayleigh mixture with $\boldsymbol{\alpha} = [1.8; 1.99]$, $\boldsymbol{\gamma} = [1; 2]$ and $\mathrm{P}[z_n = 1] = \mathrm{P}[z_n = 2] = \frac{1}{2}$. Figure 1 shows histograms of the estimated posterior density for each unknown parameter. These histograms have been generated using a single MCMC chain of 30,000 iteration (with an initial 2,000-step burn-in period). Moreover, MMSE estimates have been computed for each unknown parameter using 50 independent Monte Carlo runs (each with 1,000 iterations). The averaged MMSE estimates and the corresponding standard deviations for the different parameters are $\hat{\boldsymbol{\alpha}} = [1.81 \pm 0.020; 1.98 \pm 0.007]$ and $\hat{\boldsymbol{\gamma}} = [0.99 \pm 0.019; 1.99 \pm 0.018]$ (where each result has the form [mean $\pm$ standard deviation]). The averaged estimates of the class conditional probabilities are $\widehat{P}[z_n = 1|\boldsymbol{r}] = 0.49$ and $\widehat{P}[z_n = 1|\boldsymbol{r}] = 0.51$. We observe that all the estimates are in very good agreement with the actual values of the parameters.

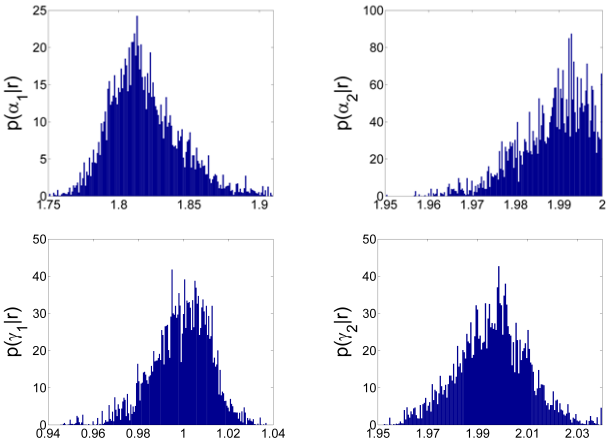


Fig. 1. Histograms of parameters generated using the proposed Gibbs sampler.

A second simulation has been performed to illustrate the method capacity to correctly label observations. This experiment has been conducted using a 3-component data set where true labels have a clear spatial arrangement. Hence, the proposed algorithm has been used to estimate the labels and retrieve this arrangement. For this experiment, the data have been generated according to a α-Rayleigh mixture model with the parameters $\alpha = [1.80; 1.90; 1.99]$, $\gamma = [50; 10; 1]$, $\mathrm{P}[z_n = 1] = \mathrm{P}[z_n = 2] = 0.455$ and $\mathrm{P}[z_n = 3] = 0.090$. Figure 2 shows the true labels $\mathbf{z}$ and their MAP estimates. The three classes can clearly be recovered even if there are misclassifications due to the intrinsic complexity of the problem.

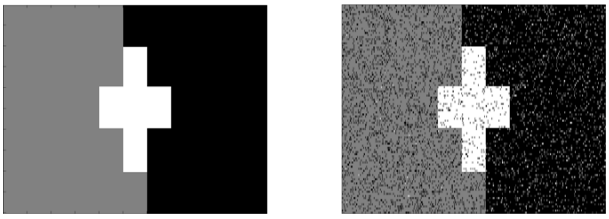


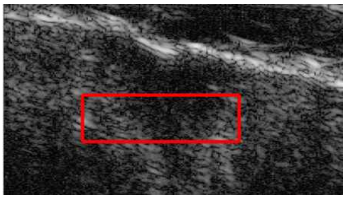
Fig. 2. True labels and MAP estimates for a 3-class mixture.

5.2. Application to real data

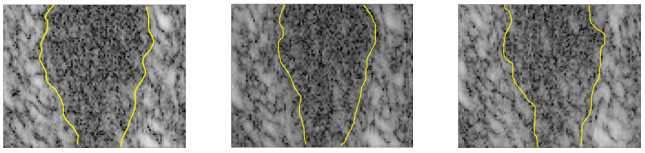
This section applies the proposed method to the detection of a skin lesion. This experiment has been conducted using a 3D high frequency ultrasound image of in-vivo skin tissues, acquired at 100MHz with a focalized 25MHz 3D probe. A skin lesion annotated by an expert is outlined by the red rectangle shown in Fig. 3. Three slices of the 3D B-mode image associated to the region of interest (ROI) are shown in Fig's 3(b), 3(c) and 3(d) where the expert annotations have been displayed in yellow.

The proposed Bayesian algorithm has been used to label each voxel as *healthy* or *lesion* tissue. Fig's 3(e), 3(f) and 3(g) show the MAP estimates for each label (*healthy* voxels are represented in white and *lesion* voxels in red). We observe that most of the MAP labels are in agreement with the expert annotations. Moreover, histograms of the estimated posterior density for each parameter are

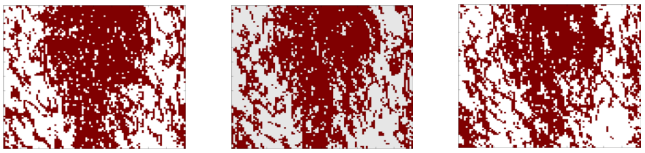
depicted in Fig. 4. These histograms indicate that the discrimination between healthy and lesion tissues can be handled by using the parameter γ of the α-Rayleigh mixture model, which significantly differs for the two classes. Conversely, the parameter α has very similar values for both classes.



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



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



(e) Labels $\boldsymbol{z}$ (slice 1) (f) Labels $\boldsymbol{z}$ (slice 2) (g) Labels $\boldsymbol{z}$ (slice 3)

Fig. 3. Log-compressed US images of skin lesion and the corresponding estimated labels (*healthy* = white, *lesion* = red).

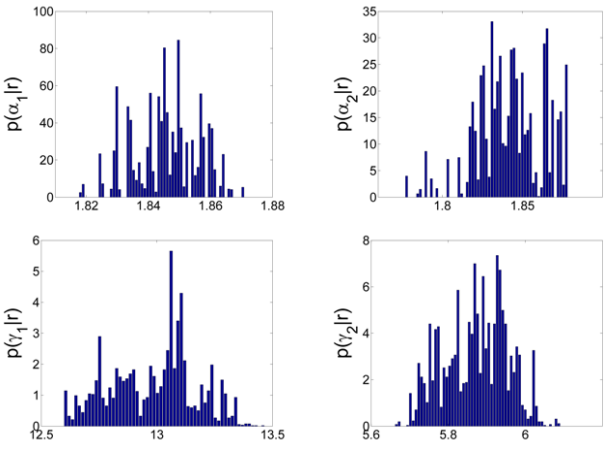


Fig. 4. [Left]: Estimated posterior densities for healthy tissue. [Right]: Estimated posterior densities for lesion tissue.

6. CONCLUDING REMARKS

A α-Rayleigh finite mixture model was proposed to represent the distribution of envelope ultrasound images backscattered from multiple tissues. An appropriate Bayesian algorithm was studied to estimate the unknown parameters of this mixture model. The proposed method does not require prior segmentation and is fully unsupervised. Moreover, it jointly estimates the mixture parameters and image labels which allow each voxel of the tissue to be classified. The application to in-vivo skin lesion in a high frequency ultrasound image is very promising. Future works include the exploitation of spatial correlations between the voxels, e.g., using Markov random fields to design an appropriate label prior, as in [4].

7. REFERENCES

[1] 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.*, May 2006.

[2] F. Destrempes, J. Meunier., 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.

[3] 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.

[4] O. Eches, N. Dobigeon, and J. Y. Tourneret, "Enhancing hyperspectral image unmixing with spatial correlations," *IEEE Trans. Geosci. Remote Sens.*, 2010, submitted. [Online]. Available: http://eches.perso.enseeiht.fr

[5] P. J. Green, "Reversible jump Markov chain Monte Carlo methods computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, Dec. 1995.

[6] M. A. Kutay, A. P. Petropulu, and C. W. Piccoli, "On modeling biomedical ultrasound RF echoes using a power-law shot-noise model," *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, vol. 48, no. 4, pp. 953–968, July 2001.

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

[8] M. L. Oelze, W. D. O'Brien, and J. F. Zachary, "Quantitative ultrasound assessment of breast cancer using a multiparameter approach," in *Proc. IEEE Ultrason. Symp.*, New York, NY, Oct. 2007.

[9] M. A. Pereyra and H. Batatia, "An alpha-stable model for ultrasound speckle statistics in skin," in *Proc. IEEE Eng. Med. Bio. Soc. (EMBC)*, Buenos Aires, Argentina, Sept. 2010.

[10] ——, "A Levy flight model for ultrasound in skin tissues," in *IEEE Proc. Ultrason. Symp.*, San Diego (CA), USA, Oct. 2010.

[11] B. I. Raju, "High frequency ultrasonic characterization of human skin in vivo," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, June 2002.

[12] S. Richardson and P. J. Green, "On Bayesian analysis of mixtures with an unknown number of components," *J. Roy. Stat. Soc. Ser. B*, vol. 59, no. 4, pp. 731–792, 1997.

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

[14] 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.

[15] J. Seabra, J. Sanches, F. Ciompi, and P. Radeva, "Ultrasonographic plaque characterization using a rayleigh mixture model," in *Proc. IEEE Symp. Med. Imag.*, Rotterdam, Netherlands, April 2010.