Restoration of Ultrasound Images Using A Hierarchical Bayesian Model with A Generalized Gaussian Prior

Ningning Zhao, Adrian Basarab, Denis Kouamé, Jean-Yves Tourneret

To cite this version:

HAL Id: hal-01399870
https://hal.archives-ouvertes.fr/hal-01399870
Submitted on 21 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://oatao.univ-toulouse.fr
Eprints ID: 15170

The contribution was presented at ICIP 2014:


Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr
RESTORATION OF ULTRASOUND IMAGES USING A HIERARCHICAL BAYESIAN MODEL WITH A GENERALIZED GAUSSIAN PRIOR

Ningning Zhao1,2, Adrian Basarab2, Denis Kouame2, Jean-Yves Tournèret1,3

1 University of Toulouse, INP/ENSEEIHT-IRIT, 2 rue Charles Camichel, BP 7122, 31071 Toulouse Cedex 7, France
2 University of Toulouse, IRIT, CNRS UMR 5505, Université Paul Sabatier, Toulouse, France
3 TéSA Laboratory, 14-16 Port Saint-Etienne, 31000 Toulouse, France
{nzhao, jean-yves.tournèret}@enseeiht.fr, {adrian.basarab, denis.kouame}@irit.fr

ABSTRACT
This paper addresses the problem of ultrasound image restoration within a Bayesian framework. The distribution of the ultrasound image is assumed to be a generalized Gaussian distribution (GGD). The main contribution of this work is to propose a hierarchical Bayesian model for estimating the GGD parameters. The Bayesian estimators associated with this model are difficult to be expressed in closed form. Thus we investigate a Markov chain Monte Carlo method which is used to generate samples asymptotically distributed according to the posterior of interest. These generated samples are finally used to compute the Bayesian estimators of the GGD parameters. The performance of the proposed Bayesian model is tested with synthetic data and compared with the performance obtained with the expectation maximization algorithm.

Index Terms— Generalized Gaussian distribution, ultrasound imaging, Bayesian inference, Gibbs sampler.

1. INTRODUCTION
Ultrasound imaging (UI) is widely used for the visualization of anatomical structures, tissue characterization and for the analysis of blood flow measurements. The popularity of UI is mainly due to its efficiency, low cost and safety for human body [1] compared to other medical imaging modalities such as Computed Tomography (CT) or Magnetic resonance imaging (MRI). However, ultrasound images are degraded by a so-called speckle noise caused by tissue inhomogeneities. Several strategies can be found in the literature to remove the speckle noise, which is very interesting for many image processing applications such as image segmentation [2,3]. However, despite the negative effect of speckle, it also contains useful textural information. Therefore, finding methods for restoring ultrasound images using the statistical properties of speckle noise is also an interesting research track [4,5].

Under the first order Born approximation which is classically assumed for soft tissues, ultrasound radio frequency (RF) images can be modeled as the 2D convolution between a blurring operator/point spread function (PSF) and the tissue reflectivity image [4,5]. The corresponding linear model can be written using a matrix-vector formulation,

\[ y = Hx + n \]  

(1)

where \( y \), \( x \), and \( n \) are obtained after lexicographical ordering the corresponding images, i.e., the RF image, the tissue reflectivity (the image to be recovered from the observed data) and the additive white Gaussian noise respectively. Note that these vectors belong to \( \mathbb{R}^{N \times 1} \) (where \( N = m \times n \) is the number of pixels of the image of size \( m \times n \)). We assume in this work that the matrix \( H \in \mathbb{R}^{N \times N} \) is a known and invariant circular matrix associated with the PSF.

The objective of image restoration is to estimate the clean original image \( x \) from the noisy measurements \( y \). Some recent works have shown that the statistical properties of the tissue reflectivity image can be described accurately by a generalized Gaussian distribution (GGD) [4,6]. This prior allows a robust ultrasound image restoration, providing the parameters of the prior model can be estimated accurately. Estimation methods based on the expectation maximization (EM) algorithm have been proposed to handle this problem [4,6]. However, the EM algorithm is known to have several shortcomings including its convergence to local minima of the likelihood yielding sometimes inaccurate estimates. This paper introduces a new hierarchical Bayesian model with a generalized Gaussian prior for the ultrasound tissue reflectivity \( x \). The posterior distribution of \( x \) is too complicated to compute closed-form expressions of the corresponding Bayesian estimators. As a consequence, we consider a Markov chain Monte Carlo (MCMC) method based on a classical Gibbs sampler which generates samples asymptotically distributed according to the posterior of interest and estimates the GGD parameters using these generated samples.

The paper is organized as follows. Section 2 introduces the proposed hierarchical model to solve the ultrasound image restoration problem. Section 3 studies a hybrid Gibbs sampler to sample the posterior distribution of the proposed model and thus to compute the Bayesian estimators of the unknown model parameters. Simulation results obtained on synthetic data are presented in Section 4. Conclusions are reported in Section 5.

2. HIERARCHICAL BAYESIAN MODEL
Bayesian inference assumes that the vector containing the unknown model parameters (denoted as \( \theta \)) is a realization of a random vector with an appropriate prior distribution. Combining this prior distribution with the likelihood of the observation model allows the posterior distribution of \( \theta \) to be computed. The unknown parameter vector \( \theta \) can then
be estimating by computing the mean or the maximum of this posterior distribution. The resulting Bayesian estima-
tors are referred to as minimum mean-squared error (MMSE) and maximum a posteriori (MAP) estimators. This section defines the likelihood of model (1) and the priors assigned to the unknown parameters of this model.

2.2.2. Noise variance

The noise \( \mathbf{n} \) in (1) is additive white Gaussian with covariance matrix \( \sigma_n^2 \mathbf{I}_{N \times N} \) (where \( \mathbf{I}_{N \times N} \) is the \( N \times N \) identity matrix). Consequently, the likelihood function of \( y \) is

\[
p(y|x, \sigma_n^2) = \frac{1}{(2\pi \sigma_n^2)^{N/2}} \exp\left( -\frac{1}{2\sigma_n^2} ||y - \mathbf{H}x||^2_2 \right) \quad (2)
\]

where \( || \cdot ||_2 \) is the usual \( \ell_2 \)-norm.

2.2. Prior Distributions

2.2.1. Reflectivity image

Following [4, 6], we assume that each pixel of the ultrasound image is distributed according to a GGD. Moreover, we assume that the ultrasound image is divided into \( K \) classes of homogeneous regions. Specifically, the pixels in one homogeneous region are independent and identically distributed according to a GGD with the same shape and scale parameters. Denote the pixels in the \( k \)th homogeneous region as

\[
x_i \sim \mathcal{GGD}(\xi_k, \gamma_k)
\]

where \( i \in \{1, \ldots, N_k\} \), \( k \in \{1, \ldots, K\} \), \( \xi_k \), \( \gamma_k \) and \( N_k \) are the shape parameter, scale parameter and total number of pixels of region \( k \) respectively. The regions can be represented by labels \( z \in \mathbb{R}^N \). More specifically, \( z_i = k \) if the corresponding pixel \( x_i \) belongs to the class \( k \). Given the pixel labels, we can obtain the following prior

\[
p(x|z) = \prod_{k=1}^{K} \prod_{i=1}^{N_k} \alpha_k \exp \left( -\frac{|x_i|^\gamma_k}{\gamma_k} \right)
= \prod_{k=1}^{K} \alpha_k^{N_k} \exp \left( -\frac{\sum_{i=1}^{N_k} |x_i|^\gamma_k}{\gamma_k} \right) \quad (3)
\]

where \( \alpha_k = \frac{1}{2^{\gamma_k/2} \Gamma(1+1/\gamma_k)} \). This paper assumes that the pixel labels are known (they are for instance obtained by a pre-segmentation or by manual annotations). As a consequence, the prior of the reflectivity image (3) is \( p(x) = p(x|z) \).

2.2.2. Noise variance

In the presence of an additive white Gaussian noise, it is very classical to assign a conjugate inverse gamma prior to the noise variance, i.e.,

\[
\sigma_n^2 \sim \mathcal{IG}(\alpha, \beta).
\]

This prior has two adjustable parameters \( \alpha, \beta \) which makes it very flexible and thus appropriate for many applications.

2.2.3. Hyperparameter Priors

The priors defined above depend on the shape parameters \( \xi = (\xi_1, \ldots, \xi_K) \) and the scale parameters \( \gamma = (\gamma_1, \ldots, \gamma_K) \) of the GGDs, which are usually referred to as hyperparameters. Following [7], we have chosen independent priors for these hyperparameters as follow

\[
p(\xi) = \prod_{k=1}^{K} p(\xi_k) \quad (5)
\]

\[
p(\gamma) = \prod_{k=1}^{K} p(\gamma_k) \quad (6)
\]

where \( 1_A \) is the indicator function on the set \( A \). [0, 3] covers all the possible values of \( \xi_k \) and \( p(\gamma_k) \) is the uninformative Jeffreys prior for \( \gamma_k \).

2.3. Joint posterior distribution

The joint posterior distribution of the unknown parameters \( x, \sigma_n^2, \xi, \gamma \) obeys the following relationship

\[
p(x, \sigma_n^2, \xi, \gamma|y) \propto p(y|x, \sigma_n^2, \xi, \gamma)p(x, \sigma_n^2, \xi, \gamma) \quad (7)
\]

where \( \propto \) means “proportional to”. Assuming the prior distributions of \( x, \sigma_n^2, \xi, \gamma \) are independent, the joint posterior of the unknown parameters can be expressed as

\[
p(x, \sigma_n^2, \xi, \gamma|y) \propto \prod_{k=1}^{K} \left( \frac{1}{\gamma_k} \right)^{1/2} \exp \left( -\frac{\beta}{\gamma_k} \right) \times \prod_{k=1}^{K} \left( \frac{1}{2^{\gamma_k/2} \Gamma(1+1/\gamma_k)} \right)^{N_k} \exp \left( -\frac{\sum_{i=1}^{N_k} |x_i|^\gamma_k}{\gamma_k} \right) \times 1_{[0,3]}(\xi_k) \times 1_{[0,\gamma]}(\gamma_k) \quad (8)
\]

The posterior (8) is clearly too complicated to compute closed-form expressions of the MMSE or MAP estimators of the unknown parameters \( x, \sigma_n^2, \xi, \gamma \). In this case, it is very classical to generate samples distributed according to this posterior by using MCMC methods. It is the objective of the next section.

3. HYBRID GIBBS SAMPLER

The principle of MCMC methods is to construct a Markov chain whose equilibrium distribution is the target distribution. The basic Gibbs sampler consists of sampling according to the conditional distributions of the target distribution. The proposed sampler is a 4-step procedure whose moves are detailed below.

3.1. Sampling the noise variance

The conditional distribution of \( \sigma_n^2|x, \xi, \gamma, y \) is

\[
p(\sigma_n^2|x, \xi, \gamma, y) \propto \left( \frac{1}{2 \pi \sigma_n^2} \right)^{N/2} \exp \left( -\frac{1}{2\sigma_n^2} ||y - \mathbf{H}x||^2_2 \right) x \left( \frac{\sigma_n^2}{\Gamma(\alpha)} \right)^{\alpha-1} \exp \left( -\frac{\beta}{\sigma_n^2} \right) \quad (9)
\]

It is the inverse gamma distribution

\[
\mathcal{IG} \left( \alpha + N/2, \beta + \frac{1}{2} ||y - \mathbf{H}x||^2_2 \right).
\]
3.2. Sampling the hyperparameters

Based on the independence assumption for the hyperparameters, the conditional distributions of the hyperparameters of the proposed GGD satisfy the following relations

\[
p(\xi_k | y, x, \sigma_n^2, \gamma, \xi_{-k}) \propto a_k^{N_k} \exp \left( -\frac{\sum_{i=1}^{N_k} |x_i|^2 \xi_k}{\gamma_k} \right) I_{[\theta, \infty]}(\xi_k) \tag{10}
\]

\[
p(\gamma_k | y, x, \xi_n^2, \gamma_{-k}) \propto a_k^{N_k} \exp \left( -\frac{\sum_{i=1}^{N_k} |x_i|^2 \xi_k}{\gamma_k} \right) \frac{1}{\gamma_k} I_{[\gamma_k, \infty]}(\gamma_k) \tag{11}
\]

where \( \theta = (\xi, \gamma), \theta_{-k} = (\theta_1, ..., \theta_{k-1}, \theta_{k+1}, ..., \theta_K) \). The distribution (10) is not easy to sample directly. Thus, we propose to consider zero mean Gaussian random walks (whose variances are adjusted a priori to obtain a reasonable acceptance ratio) as the candidates and then to accept or reject the proposal according to the classical MH acceptance ratio (8). The distribution (11) is an inverse gamma distribution that is easy to sample.

3.3. Sampling the reflectivity image

The conditional distribution of the image we want to estimate is defined as follows

\[
p(x | y, \sigma_n^2, \xi, \sigma_\xi^2) \propto \exp \left( -\frac{1}{2\sigma_n^2} ||y - Hx||_2^2 - \frac{1}{2} \sum_{k=1}^{K} \sum_{i=1}^{N_k} |x_i|^2 \xi_k \right) \tag{12}
\]

Sampling according to (12) is the critical point of the proposed algorithm. Due to the high dimensionality of \( x \), classical Gibbs or MH moves are inefficient. Thus, we propose an efficient sampling strategy based on Hamiltonian Monte Carlo/Hybrid Monte Carlo (HMC). This method makes use of Hamiltonian dynamics and gradient information in order to fasten the convergence of the sampler. Some elements allowing the principles of HMC methods to be understood are provided in the next section. The reader is invited to consult [9] for more details.

3.3.1. Hamiltonian Monte Carlo Algorithm

The main idea of HMC is to introduce a vector of momentum variables \( p \in \mathbb{R}^N \) and to sample the pair \((x, p)\) instead of just sampling \( x \). The conditional distribution of \((x, p)\) is

\[
p(x, p | y, \sigma_n^2, \xi, \sigma_\xi^2, \gamma) = p(x | y, \sigma_n^2, \xi, \gamma)p(p).
\]

The Hamiltonian of the system is defined as

\[H(x, p) = -\log p(x, p | y, \sigma_n^2, \xi, \gamma) = U(x) + K(p)\]

where \( K(p) \) and \( U(x) \) are the kinetic and potential energies

\[K(p) = \frac{1}{2}p^T p \quad \text{and} \quad U(x) = -\log[p(x | y, \sigma_n^2, \xi, \gamma)].\]

At the iteration \( t \), the HMC consists of two steps: 1) generate a candidate pair \((p^{(t)}^*, x^{(t)}^*)\) from the current state \((p^{(t)}, x^{(t)})\) using a discretization method, such as Leapfrog and Euler methods; 2) accept or reject the candidate with the probability

\[
\rho = \min\{ \exp[\mathcal{H}(x^{(t)}) - \mathcal{H}(x^{(t)}^*)], 1 \}. \tag{13}
\]

In our experiments, we have implemented the Leapfrog method defined below

\[
p_1(t + \epsilon/2) = p_1(t) - \epsilon \frac{\partial U}{\partial x_i}[x(t)]
\]

\[
x_i(t + \epsilon) = x_i(t) + \epsilon p_i(t + \epsilon/2)
\]

\[
p_1(t + \epsilon) = p_1(t + \epsilon/2) - \epsilon \frac{\partial U}{\partial x_i}[x(t + \epsilon)]
\]

where \( \epsilon \) is a so-called stepsize and \( L \) is the number of Leapfrog steps per trajectory. Note that we have not explored other discretization alternatives since we have obtained interesting results with the Leapfrog scheme. Compared to other

**Algorithm 1: Hamiltonian Monte Carlo Algorithm**

/* Initialization of the \( t \)th iteration */
1. \( p^{(1)} = N(0, I) \)
2. \( x^{(1)} = y \)
3. \( p^{(1)} \leftarrow p^{(1)} \)
4. \( x^{(1)} \leftarrow x^{(1)} \)

/* Leapfrog Method */
5. for \( i = 1 : L \) do
6. \( \text{Set } p^{(i)} = p^{(i)} - \frac{\epsilon}{2} \frac{\partial U}{\partial x_i}[x^{(i)}] \)
7. \( \text{Set } x^{(i)} = x^{(i)} + \epsilon p^{(i)} \)
8. \( \text{Set } p^{(i)} = p^{(i)} - \frac{\epsilon}{2} \frac{\partial U}{\partial x_i}[x^{(i)}] \)
9. end

/* Accept/Reject Procedure */
10. Compute \( \rho \) with (13)
11. Set \( (x^{(t+1)}, p^{(t+1)}) = (x^{(t)}, p^{(t)}) \) with probability \( \rho \)
12. Else set \( (x^{(t+1)}, p^{(t+1)}) = (x^{(t)}, p^{(t)}) \)

MCMC algorithms, the HMC method has the noticeable advantage to generate a candidate \( x \) in three steps (detailed in Algorithm 2) and to accept or reject this candidate using the acceptance ratio (13).

3.3.2. Tuning the parameters \( \epsilon \) and \( L \)

The performance of the HMC algorithm mainly depends on the values of parameters \( \epsilon \) (stepsize) and \( L \) (number of Leapfrog steps). Fortunately, these two parameters can be tuned independently in most applications [9]. It is recommended to select a random number of Leapfrog steps \( L \) to avoid possible periodic trajectories [9]. In our algorithm, \( L \) is sampled uniformly in the interval [50, 70] [9]. The leapfrog stepsize \( \epsilon \) has been adjusted by cross validation in order to ensure a reasonable acceptance rate.

3.4. Parameter estimation

The Bayesian estimators can be computed using the samples generated by the proposed MCMC method. For instance, the
MMSE estimator of $x$ can be approximated by an average of the generated samples $\hat{x}_{\text{MMSE}} \approx \frac{1}{N_{\text{MC}}-N_{\text{bi}}} \sum_{t=0}^{N_{\text{MC}}} x^{(t)}$, where $N_{\text{bi}}$ is the number of burn-in iterations required to reach the sampler convergence and $N_{\text{MC}}$ is the total number of Monte Carlo iterations.

4. SIMULATIONS

The objective of this section is to evaluate the performance of the proposed hierarchical Bayesian model (based on a GGD prior) for the restoration of synthetic ultrasound data. A comparison with the EM algorithm developed in [4, 6] is also presented. The ultrasound data was simulated by 2D convolution between a known PSF and the reflectivity image as in [5]. The PSF was simulated with Field II [10]. The amplitudes of the scatters in the reflectivity image are random variables distributed according to GGDs. Specifically, $x_A \sim (\xi=1.5, \gamma=38.7120)$ for disk A, $x_B \sim (\xi=0.9, \gamma=4.8079)$ for disk B, $x_C \sim (\xi=2, \gamma=188.9059)$ for disk C, $x_D \sim (\xi=1, \gamma=6.8029)$ for disk D and $x_E \sim (\xi=0.5, \gamma=0.3025)$ for the background.

The resulting restored image is displayed in Fig. 1(d), illustrating the good deconvolution performance of the proposed algorithm. In order to provide more quantitative results, we have considered the results of Part E in Fig. 1 and have compared the performance of the proposed algorithm with that of the EM approach studied in [4, 6]1. Note that the size of Part E is $50 \times 50$ and that the noise power corresponds to a blurred signal to noise ratio BSNR= 40dB (see e.g. [11] for a definition of BSNR). The histograms of the parameters generated by the proposed hybrid Gibbs sampler for the pixels of Part E are displayed in Fig. 2. These histograms are clearly in good agreement with the true values of the parameters indicated by the vertical lines, which confirms the good performance of the proposed method. The improvement in SNR (ISNR) (see e.g. in [11]) for the proposed method is shown in Fig. 2 versus the number of iterations of the proposed algorithm. Table 1 shows more quantitative results in terms of ISNR, normalized root mean square error (NRMSE) [12], structural similarity (SSIM) [13] and peak SNR (PSNR) [14] which are compared with the results obtained with the EM algorithm of [4, 6]. And the proposed strategy gives better results than the EM algorithm. Table 2 compares some estimators of the GGD parameters obtained with the EM and the proposed algorithm. These results confirm the good performance of the proposed method.

![Image](a) Tissue reflectivity  
(b) Ultrasound RF image  
(c) Ultrasound B-mode image  
(d) Restored image

Fig. 1. (a) Tissue reflectivity map $x$. (b) Ultrasound RF image $y$. (c) Ultrasound B-mode image (i.e., log compressed envelop of $y$). (d) Restored image.

![Image](Fig. 2. The histograms of the sampled marginal posterior distributions of $\sigma^2_n$ (top left), the hyperparameters $\xi$ (top right) and $\sigma^2_x$ (bottom left). Their corresponding ground truth values (vertical lines). In bottom right, ISNR (dB) versus Monte Carlo iterations (after the burn-in period) for part E.)

<table>
<thead>
<tr>
<th>Methods</th>
<th>ISNR(dB)</th>
<th>NRMSE(dB)</th>
<th>SSIM</th>
<th>PSNR(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>9.549</td>
<td>0.3429</td>
<td>0.8355</td>
<td>30.4740</td>
</tr>
<tr>
<td>Proposed</td>
<td>1.111e-05</td>
<td>0.6388</td>
<td>0.4202</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. True values of the parameters and their estimations.

<table>
<thead>
<tr>
<th>Methods</th>
<th>$\sigma^2_{n}$</th>
<th>$\xi_{E}$</th>
<th>$\gamma_{E}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EM</td>
<td>2.6e-3</td>
<td>0.5397</td>
<td>-</td>
</tr>
<tr>
<td>true value</td>
<td>1.1434e-05</td>
<td>0.5</td>
<td>0.3021</td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

This paper proposed a Bayesian model with a generalized Gaussian prior for the deconvolution of ultrasound images. An MCMC method was proposed to sample according to the posterior of this model and to estimate the unknown model parameters. The results obtained on synthetic images are very promising. However, the limitation of this work is that the pixel labels of image were supposed to be known. Future work includes the study of a segmentation algorithm for estimating of the pixels labels and the application of the proposed algorithm to real data.

1The authors would like to thank M. Alessandrini for sharing his codes with the authors of this paper.
6. REFERENCES


