NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET

Bayesian 2D Deconvolution: A Model for Diffuse Ultrasound Scattering

by

Oddvar Husby, Torgrim Lie, Thomas Langø, Jørn Hokland and Håvard Rue

PREPRINT
STATISTICS NO. 20/1999

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL http://www.math.ntnu.no/preprint/statistics/1999/S20-1999.ps
Oddvar Husby has homepage: http://www.math.ntnu.no/∼okh
E-mail: okh@math.ntnu.no
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

Bayesian 2D Deconvolution: A Model for Diffuse Ultrasound Scattering

Oddvar Husby, Torgrim Lie, Thomas Langø,
Jørn Hokland and Håvard Rue

Abstract

Observed medical ultrasound images are degraded representations of the true acoustic tissue reflectance. The degradation is due to blur and speckle, and significantly reduces the diagnostic value of the images. In order to remove both blur and speckle we have developed a new statistical model for diffuse scattering in 2D ultrasound radio-frequency images, incorporating both spatial smoothness constraints and a physical model for diffuse scattering. The modeling approach is Bayesian in nature, and we use Markov chain Monte Carlo methods to obtain the restorations. The results from restorations of some real and simulated radio-frequency ultrasound images are presented, and compared with results produced by Wiener filtering.

Keywords

Medical ultrasound, restoration, Markov random fields, diffuse scattering, Markov chain Monte Carlo

I. Introduction

Ultrasound B-scan images are degraded representations of the true acoustical reflectors in the imaged tissue. The degradation introduces image artifacts (speckle) and limits the spatial resolution, hence reducing the diagnostic value of the images. The dominant degradation effects are blur and speckle due to what may be approximated as a convolution of the true reflectance field r with the imaging system point spread function h, with the addition of noise introduced in the image formation process. We assume the point spread function to be spatially invariant[1], [2], and thus the observed signal y can be modeled as

$$y_{i,j} = \{h * r\}_{i,j} + \eta_{i,j} = \sum_{k,l} h_{k,l} r_{i-k,j-l} + \eta_{i,j}, \quad (i,j) \in \mathcal{I} \tag{1}$$

where $\eta_{i,j}$ are independent Gaussian variables with mean zero and common variance τ^2, and $\mathcal{I}$ is a rectangular lattice with $n_r \times n_l = n$ sites. The indices i and j represent radial and lateral location, respectively. Note that the assumption of spatially invariant point spread function is made out of computational convenience only, since our focus is on the modeling of the image formation process. We study the effect of this assumption in [3], and refer to e.g. [4], [5] for other discussions of the effect of the point spread function.

Due to the limited bandwidth of the point spread function and the presence of additive noise, perfect deconvolution is impossible. As an alternative, statistical restoration of ultrasound images attempts to remove the effect of the degradation by estimating the true reflectance image r. A simple and commonly used technique is Wiener filtering [6], [1]; other methods include multiscale wavelet analysis [7], but even with these methods the restored images suffer from blur and speckle.

We propose a new method for deconvolution of ultrasound images, based on Markov random field models [8] and Bayesian statistical methods [9], [10]. These methods have become increasingly popular over the last decade or two, and represent a general methodology for restoration of indirectly observed measurements. The most important feature is the ability to model the randomness and uncertainty inherent in the system. Specifically, in dealing with ultrasound the methodology allows us to use stochastic models of the imaged tissue and the image formation process, thus incorporating *a priori* knowledge and measurement error in a way not possible in deterministic restoration methods. Also, we are able to quantitatively assess the errors in our restorations, producing a *distribution* over the image space rather than a single filter-image restoration.

O. Husby and H. Rue are with the Department of Mathematical Sciences, NTNU, N-7491 Trondheim, Norway. E-mail: Oddvar.Husby@math.ntnu.no

T. Lie and T. Langø are with SINTEF Unimed, Ultrasound, 7465 Trondheim, Norway

J. Hokland is with the Department of Informatics, NTNU, 7491 Trondheim, Norway

Our Markov random field model is based on the standard model for diffuse scattering [11], [12], where a point reflector within one resolution cell is modeled to be Gaussian with mean zero and variance depending on the tissue type within the cell. Our idea is to let a spatially varying variance field represent the *underlying structures* of the imaged region; based on the assumption that this field is the *sole* descriptor needed to obtain faithful image restorations. The argument for the scattering model in [12] goes as follows: Assuming that the concentration of reflectors is large relative to the wavelength of the pulse, the reflectance $r_{i,j}$ at each location (i,j) is a Gaussian random variable with mean zero and variance $\sigma^2_{i,j}$. The reflectances at different points are assumed to be uncorrelated, and thus the reflectances has the conditional distribution (i.e. the distribution of $r_{i,j}$ given that the value of $\sigma^2_{i,j}$ is known)

$$p(r_{i,j} \mid \sigma^2_{i,j}) = \frac{1}{\sqrt{2\pi\sigma^2_{i,j}}} \exp\left\{ -\frac{r^2_{i,j}}{2\sigma^2_{i,j}} \right\}, \ \forall\, (i,j) \in \mathcal{I}. \tag{2}$$

We see that the estimation of the variance field is essential; given the variance field, the reflectance can be simulated by drawing from the distribution (2).

The restoration can now be formulated as a Bayesian estimation problem. We define an appropriate prior distribution $p(\sigma^2)$ for the variance field, trying to model the spatial properties inherent in the imaged regions. Using Bayes theorem, the prior distribution is then combined with the likelihood model $p(y \mid \sigma^2)$ to obtain the posterior distribution $p(\sigma^2 \mid y, h)$ from which inferences can be made:

$$p(\sigma^2 \mid y, h) \propto p(y \mid \sigma^2, h) p(\sigma^2). \tag{3}$$

We define the prior model in section II, selecting a distribution from the family suggested in [13], where it was used for tomographic reconstruction. These models have the property that they are able to recover discontinuities in otherwise smooth images. This property is essential in our application, as discontinuities may represent interfaces between regions of different tissue types. Following the definition of the prior we compute the full posterior distribution. The distribution is analytically intractable; hence we have to rely on iterative stochastic techniques to make inferences. Markov chain Monte Carlo methods [9], [14], [15] are well suited for sampling Markov random field models, but we argue that when the model is formulated in terms of σ^2 only, these techniques can be very inefficient. However, in section III we show that this problem can be solved by augmenting the model with the reflectance field r. We show that y and σ^2 are conditionally independent given r, and thus that r can be viewed as an auxiliary variable [16], [17] decoupling y and σ^2. When sampling from the joint distribution of r and σ^2 this fact can be utilized to construct an algorithm with better computational properties.

A related approach can be found in [18], who use a discrete Markov model for the variance field. However, their discrete model made the method unstable. Furthermore, it is unclear how well a discrete model describes the underlying structure of e.g. human tissue, and also how well it can account for the radial reduction in intensity seen in ultrasound images. Our main contribution is the formulation of the continuous variance field, which is a better descriptor of the imaged regions, and which can model radial reduction in intensity. The new model is also considerably easier to work with, and from a computational point of view the continuous representation is preferable, making sampling and estimation easier, faster and more robust.

We conclude the paper with examples and a brief discussion in sections IV and V.

II. Bayesian Model Formulation

A. Introduction

With a slight abuse of notation we will use $p(\cdot)$ to represent both the distribution and the density of the random variables. Recall that the images are defined on the rectangular lattice $\mathcal{I}$, using the polar coordinate representation of B-scan ultrasound images. All images are defined on the domain $\mathcal{I}$, letting e.g. $\sigma^2 = \left(\sigma^2_{1,1}, \dots, \sigma^2_{n_r,n_l}\right)^T$, where n_r and n_l represent the number of samples in the radial and lateral directions, respectively. For any subset $\mathcal{A} \subset \mathcal{I}$ we write $r_{\mathcal{A}} = \{r_{i,j}\,;\,(i,j) \in \mathcal{A}\}$ and $r_{-\mathcal{A}} = \{r_{i,j}\,;\,(i,j) \in \mathcal{I} \setminus \mathcal{A}\}$, and write $p(r_{\mathcal{A}})$ and $p(r_{-\mathcal{A}})$ for the corresponding densities. Our approach is based on Markov random field (MRF) models. One of the key features of Markov random fields is that each site (i,j) has associated with it a neighborhood $\partial(i,j)$, such that the conditional probability of the field

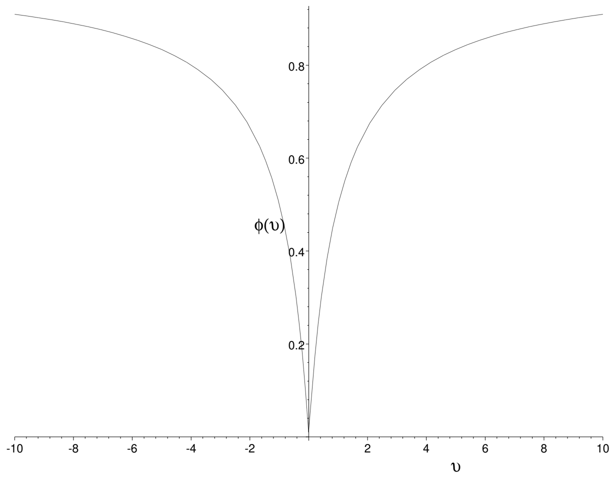

Fig. 1. Plot of the interaction function $\phi(u)$ with $\delta = 1$ and $\gamma = 1$.

at the site (i, j), given the values at all other sites, only depend on the values of the field in $\partial(i, j)$ [8]. We write $(i, j) \sim (k, l)$ if the sites (i, j) and (k, l) are neighbors.

B. Prior distribution for the variance field

The crucial part of our model is the formulation of the prior distribution for the variance field σ^2. We let the prior be of the form

$$p(\sigma^2) \propto \exp\left\{-\beta \sum_{i,j\sim k,l} w_{ij,kl}\,\phi\left(\ln\left(\sigma^2_{i,j}/\sigma^2_{k,l}\right)\right) - \sum_{i,j} \psi(\sigma^2_{i,j})\right\},$$

where $\phi(\cdot)$ is a spatially structured interaction term and $\psi(\cdot)$ is an unstructured level term. The choice of the prior model is justified from physical considerations about the scattering medium, and is explained below.

In the diffuse scattering model the variance depends on the acoustical properties of the scattering medium [11], and it is thus reasonable to assume the following:

- Variances tend to be approximately constant in regions of homogeneous tissue.
- Abrupt changes in variance may occur at interfaces between different tissue regions, as the scattering intensity varies between tissue types.

In other words, we believe the field to be piecewise planar with each subregion corresponding to the different tissue types in the imaged region. Hence prior distributions with convex interaction functions will be unsuited for our purpose, since they tend to oversmooth images, favoring gradual changes in intensity. We will instead use an interaction function of the form

$$\phi(u) = \frac{|\frac{u}{\delta}|^\gamma}{1 + |\frac{u}{\delta}|^\gamma}, \tag{4}$$

as suggested in [13]; this model is known to provide for recovery of discontinuities in blurred and noisy images. The noninterpolating property of the model is due to its concavity and finite limit as $u \to \infty$, favoring large intensity jumps instead of gradual changes; see figure 1. We refer to [19] and [20] for a further discussion of the properties of this class of models.

In addition to the spatially structured model (4), the unstructured term is defined as a vague inverse Gamma distribution; and can be seen as a prior on the level of the variances. Combining the two terms, we get the following prior density:

$$\begin{aligned} p(\sigma^2) &\propto \exp\left\{-\beta \sum_{i,j\sim k,l} w_{ij,kl}\,\phi\left(\ln\left(\sigma^2_{i,j}/\sigma^2_{k,l}\right)\right) - \sum_{i,j}\left(\epsilon/\sigma^2_{i,j} + (\epsilon+1)\ln\sigma^2_{i,j}\right)\right\} \\ &= \exp\left\{-\Phi(\sigma^2) - \Psi(\sigma^2)\right\}, \end{aligned} \tag{5}$$

where β and ϵ are positive parameters, and $w_{ij,kl}$ is a weight depending on the neighborhood system. In the experiments we use a second order neighborhood system with $w_{ij,kl} = 1$ for orthogonal nearest neighbors, and $w_{ij,kl} = 1/\sqrt{2}$ for diagonal nearest neighbors. The parameter β determines the strength of the interaction, while δ in eq. (4) determines the size of the intensity jumps allowed. The parameter γ is set equal to 1 in all experiments, since for $\gamma \leq 1$ sharp boundaries are favored over gradual transitions, while larger values of γ permit more variations within homogeneous regions.

C. Posterior distribution

As noted in section I, it is possible to formulate the model in terms of the variance field σ^2 only. Given an estimate $\hat{\sigma}^2$, the reflectance field r can be simulated readily using eq. (2). However, in appendix A we show that the posterior density $p(\sigma^2 \,|\, y)$ has poor sampling properties, leading to computational problems. This deficiency is amended by considering the augmented model (σ^2, r) as follows.

C.1 Model augmentation

The idea behind model augmentation is to introduce auxiliary variables that allow one to construct Markov chains that are easier to simulate. The variable z of interest is augmented by a variable u, and to generate realizations from $p(z)$ one specifies the conditional distribution $p(u \,|\, z)$ and builds $p(z, u) = p(u \,|\, z)p(z)$, which is then sampled from. The idea is to specify u in such a way that the joint sampling of (z, u) is in some sense easier than the sampling of z alone. Note that $p(z, u)$ maintains the marginal distribution $p(z)$.

In our setting we let r be the auxiliary variable. The augmented model is then $\left(\sigma^2, r\right)$ with distribution

$$p(\sigma^2, r \,|\, y, h) \propto p(y \,|\, r, h)p(r, \sigma^2) \propto p(y \,|\, r, h)p(r \,|\, \sigma^2)p(\sigma^2). \tag{6}$$

Note that σ^2 and y are conditionally independent given r, thus r *decouples* σ^2 and y, making the variance field σ^2 conditionally independent of the blur. This is important because the blur causes long-range dependency, making unconditional sampling of σ^2 very difficult.

C.2 Posterior density

We now proceed to find the joint density for σ^2 and r. According to eq. (1), the density for the observed image y can be written as

$$p(y \,|\, r, h) \propto \exp\left\{ -\frac{1}{2\tau^2} \sum_{i,j} \left(y_{i,j} - \{h * r\}_{i,j} \right)^2 \right\}, \tag{7}$$

where the convolution between h and r can be considered a constant shifting of the expectation of y.

By inserting the expression for $p(\sigma^2)$, $p(r \,|\, \sigma^2)$, and $p(y \,|\, r, h)$ in eq. (6), the joint density for σ^2 and r can be written as

$$\begin{aligned} p(\sigma^2, r \,|\, y, h) &\propto p(y \,|\, r, h)p(r \,|\, \sigma^2)p(\sigma^2) \\ &\propto \exp\left\{ -\sum_{i,j} \left(\frac{1}{2\tau^2} \left(y_{i,j} - \{h * r\}_{i,j} \right)^2 + \frac{r_{i,j}^2 + 2\epsilon}{2\sigma_{i,j}^2} + \left(\epsilon + \frac{3}{2} \right) \ln \sigma_{i,j}^2 \right) \right. \\ &\qquad \left. -\beta \sum_{i,j \sim k,l} w_{ij,kl} \phi \left(\ln \left(\sigma_{i,j}^2 / \sigma_{k,l}^2 \right) \right) \right\}. \end{aligned} \tag{8}$$

We will later show that our sampling algorithm will use the conditional distribution of $r_{i,j}$ and $\sigma_{i,j}^2$ at single pixels (i,j), and from eq. (8) we arrive at

$$p(\sigma^2_{i,j}\,|\,\sigma^2{}_{-(i,j)},r) \propto \exp\left\{-\left(\epsilon+\frac{r^2_{i,j}}{2}\right)\frac{1}{\sigma^2_{i,j}}-\left(\epsilon+\frac{3}{2}\right)\ln\sigma^2_{i,j}-\beta\sum_{(k,l)\in\partial(i,j)} w_{ij,kl}\,\phi\left(\ln\sigma^2_{i,j}/\sigma^2_{k,l}\right)\right\} \tag{9}$$

$$p(r_{i,j}\,|\,r_{-(i,j)},\sigma^2,y,h) \propto \exp\left\{-\frac{1}{2\tau^2}\sum_{m,n}\left(y_{m,n}-\sum_{(k,l)\neq(m-i,n-j)} h_{k,l}r_{m-k,n-l}-h_{m-i,n-j}r_{i,j}\right)^2-\frac{r^2_{i,j}}{2\sigma^2_{i,j}}\right\}. \tag{10}$$

By performing the multiplication of the quadratic parenthesis and rearranging the terms, eq. (10) can be written as:

$$\begin{aligned} p(r_{i,j}\,|\,r_{-(i,j)},\sigma^2,y,h) &\propto \exp\left\{-r^2_{i,j}\left(\frac{1}{2\tau^2}\sum_{m,n}h^2_{m-i,n-j}+\frac{1}{2\sigma^2_{i,j}}\right)+2r_{i,j}\left(\frac{1}{2\tau^2}\sum_{m,n}\left(h_{m-i,n-j}\Upsilon_{m,n}\right)\right)\right\} \\ &\propto \exp\left\{-r^2_{i,j}g_1(\sigma^2,h,\tau)+2r_{i,j}g_2(r_{-(i,j)},y,h,\tau)\right\}, \end{aligned} \tag{11}$$

where

$$\Upsilon_{m,n} = y_{m,n}-\sum_{(k,l)\neq(m-i,n-j)} h_{k,l}r_{m-k,n-l} = y_{m,n}-\sum_{k,l}h_{k,l}r_{m-k,n-l}+h_{m-i,n-j}r_{i,j}. \tag{12}$$

Concerning implementation, note that we can store the values of $\Upsilon_{m,n}$ in a two-dimensional table, reducing the the double summation in eq. (10) to a single sum over the values of $\Upsilon_{m,n}$. For every pixel updated, we have to update as many values in the table as there are elements in the point spread function h. Thus we reduce the effective neighborhood size of r from the support of h convolved with h to only two times the support of h.

From eq. (11) one can recognize that the conditional distribution for $r_{i,j}$ given $y, r_{-(i,j)}$ and σ^2 to be a Gaussian distribution with mean

$$\mu_{i,j} = \frac{g_2}{g_1} = \frac{\frac{1}{\tau^2}\sum_{m,n}h_{m-i,n-j}\Upsilon_{m,n}}{\frac{1}{\tau^2}\sum_{m,n}h^2_{m-i,n-j}+\frac{1}{\sigma^2_{i,j}}}, \tag{13}$$

and variance

$$\kappa^2_{i,j} = \frac{1}{2g_1} = \left[\frac{1}{\tau^2}\sum_{m,n}h^2_{m-i,n-j}+\frac{1}{\sigma^2_{i,j}}\right]^{-1}, \tag{14}$$

and thus the conditional sampling of $r_{i,j}$ is straightforward using one of the standard algorithms for sampling Gaussian random variables.

III. The Sampling Algorithm

A. Metropolis-Hastings algorithm

Given the above model we want to use the *posterior means* as estimators of the variance and reflectance fields. The densities (8) are analytically intractable, hence we have to rely on Markov chain Monte Carlo methods as follows: We construct an ergodic Markov chain $\sigma^2(0), r(0), \sigma^2(1), r(1), \ldots$ converging in distribution to $p(\sigma^2, r\,|\,y,h)$. This is done as follows: For each $k = 1, 2, \ldots$ we generate $\sigma^2(k)$ from

$r(k-1)$ using $p(\sigma^2(k)\,|\,r(k-1))$, and then we generate $r(k)$ from $\sigma^2(k)$ using $p(r(k)\,|\,\sigma^2(k),y,h)$. After a burn-in of t_0 iterations an estimate of the posterior means can be found as

$$\widehat{\mathrm{E}\left(\sigma^2\,|\,r\right)} = \frac{1}{t}\sum_{i=t_0+1}^{t_0+t}\sigma^2(i) \qquad \widehat{\mathrm{E}\left(r\,|\,\sigma^2,y,h\right)} = \frac{1}{t}\sum_{i=t_0+1}^{t_0+t} r(i). \tag{15}$$

Several methods exist for constructing the Markov chain $\sigma^2(0), r(0), \sigma^2(1), r(1), \ldots$; we refer to [21] for a overview of the theory and methodology related to Markov chain Monte Carlo methods. Among the most common methods is the Metropolis-Hastings algorithm [14], which is the one we have used in the experiments. Assuming we want to sample from a distribution $p(z)$, the algorithm is as follows: If the current state of the Markov chain is z, a new state z' is drawn from a proposal distribution with density $q(z,z')$. The new state is then accepted with probability

$$\alpha\left(z,z'\right) = \min\left\{1, \frac{p(z')q(z',z)}{p(z)q(z,z')}\right\}, \tag{16}$$

or otherwise the old state is retained. This algorithm is usually based on a succession of random local changes, since global changes tend to have small probability of being accepted. Usually, updating is most conveniently done at one site at the time, requiring access to the full conditionals $p(z_{i,j}\,|\,z_{-(i,j)})$ and relying on the Markov property $p(z_{i,j}\,|\,z_{-(i,j)}) = p(z_{i,j}\,|\,z_{\partial(i,j)})$ to do efficient updating. The attractiveness of this approach comes from the fact that $p(z_{i,j}\,|\,z_{\partial(i,j)})$ usually has a small support, and hence can be computed efficiently.

We will follow this approach, sampling from the conditional densities given in eqs. (9) and (10). This sampling scheme is efficient for the following reasons.

- $p(\sigma^2_{i,j}\,|\,\sigma^2_{-(i,j)}, r)$ is a distribution with small support: The neighborhood is of the same size as specified in the prior (5). Because of this local Markov property, sampling is eased. We again stress that *given* r, the variance field σ^2 is independent of the point spread function h.
- $p(r_{i,j}\,|\,r_{-(i,j)}, \sigma^2, y, h)$ is a Gaussian distribution, which is straightforward to sample from.

Details of the sampling are given in the next section.

B. Sampling algorithm

We now proceed to describe some details related to the sampling scheme given in section III-A; this part may be skipped at a first reading.

B.1 Sampling the variance field

The choice of proposal distribution $q(\cdot)$ poses a small problem. Since variance is a scale parameter, we chose to propose an new state by scaling the existing state with a random number $u \sim U[1/2, 2]$. However, when working with the model, we found that the restorations often contained outlying pixel values in otherwise smooth regions. While not discussed in the original paper by [19], this is a feature of the model, as shown by [22]. The reason is as follows: Let $r^0_{i,j}$ be the true pixel value at site (i,j), and assume that $\tilde{r}_{i,j}$ is the marginal minimizer of (7). Then [22] show that the difference $\tilde{r}_{i,j} - r^0_{i,j}$ is Gaussian distributed with variance $(\frac{1}{\tau^2}\sum_{k,l} h^2_{k,l})^{-1}$. This variance can get extremely large when the blurring is severe, or when there is much noise. Then because of the concavity of the roughness penalty $\phi(\cdot)$, outlying pixel values are allowed to occur. [19] does not encounter this problem because they use a modified Gibbs sampler supported on a interval around the current values of the state and its neighbors. The algorithm is formalized in [23], restricting the image space to those images in which no pixel values differs from its neighbors by more than a threshold γ_0. However, as noted in [22], this restriction can be removed by embedding the modified Gibbs sampler in a Hastings algorithm.

In the present context, with σ^2 taking values on the whole of $\mathbb{R}^+$, Gibbs sampling is costly to implement. Since the unsatisfactory reconstructions represent local, coordinate-wise minima, a better approach would be to design the algorithm to effectively escape such local minima. We have chosen a new proposal in which, letting $\tilde{\sigma}^2_{i,j}$ be the median of the neighbors of the current state, a new state is proposed by scaling $\tilde{\sigma}^2_{i,j}$ with a random number $u \sim U[1/2, 2]$. Thus if the current state is an outlying pixel value, large

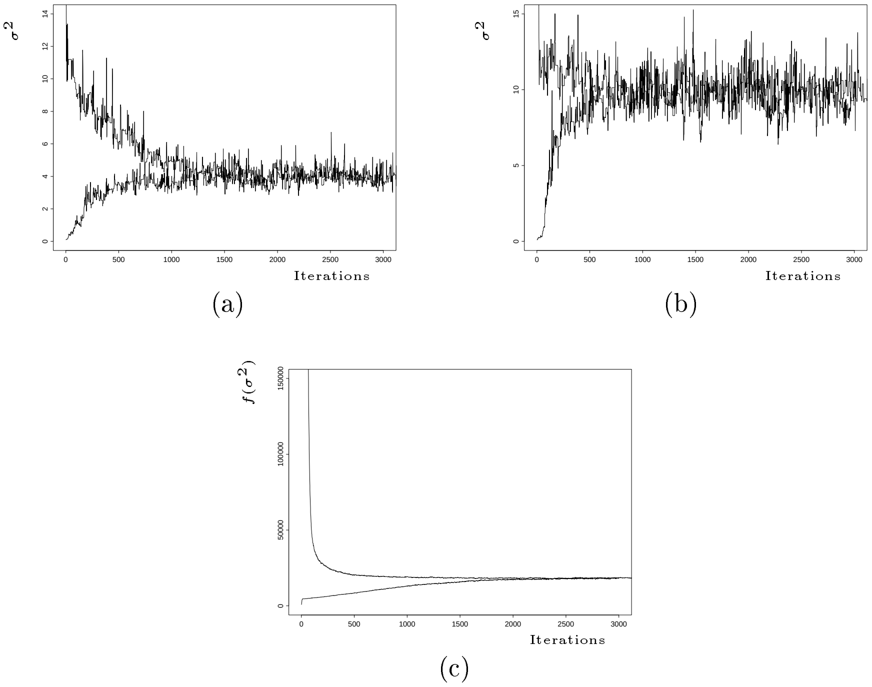

Fig. 2. Traceplots of two Markov chains started in different initial states. (a) and (b) show plots of the variance at two different sites, while (c) shows traceplot of the functional f.

jumps can be made. The choice of the median is important; choosing the mean could lead to undesired smoothing at region boundaries.

The solution suggested in [22], where the parameter β is increased to enforce a higher degree of smoothness, would also partly eliminate the problem, but at the cost of slower convergence.

B.2 Sampling the reflectance field

As customary when sampling from Gaussian distributions, we choose a Gaussian distribution centered at the current state as a proposal distribution for $r_{i,j}$. Following [24], the standard deviation of the proposal distribution is scaled so that the acceptance ratio (16) is about 0.25. This means that we accept about one out of four proposed moves.

C. Diagnosing convergence

When designing algorithms, assessing the convergence of the Markov chain is an important issue. Quantitative convergence rates are in most cases hard to find, but there exists a number of convergence diagnostics [25], [26], [27] which can be used as a guide to the convergence of the chain. In this paper we have assessed the rate of convergence by monitoring the functional

$$f(\sigma^2) = \beta \sum_{i,j \sim k,l} w_{ij,kl}\, \phi(\ln \sigma^2_{i,j}/\sigma^2_{k,l}),$$

using as burn-in the approximate number of iterations needed until all appear to have converged. To further verify that the algorithm produces consistent results, we have started sampling with over-dispersed initial states. Figure 2 shows the result from such an experiment. We start two different runs in initial σ^2-fields with constant levels of 10 an 0.1, respectively. Figures 2(a) and 2(b) shows the trace of the runs at two different locations, while 2(c) show traces for the functional f. Clearly, this is a subjective method, but the plots do at least indicate that the estimates produced by the algorithm are independent of the initial values used.

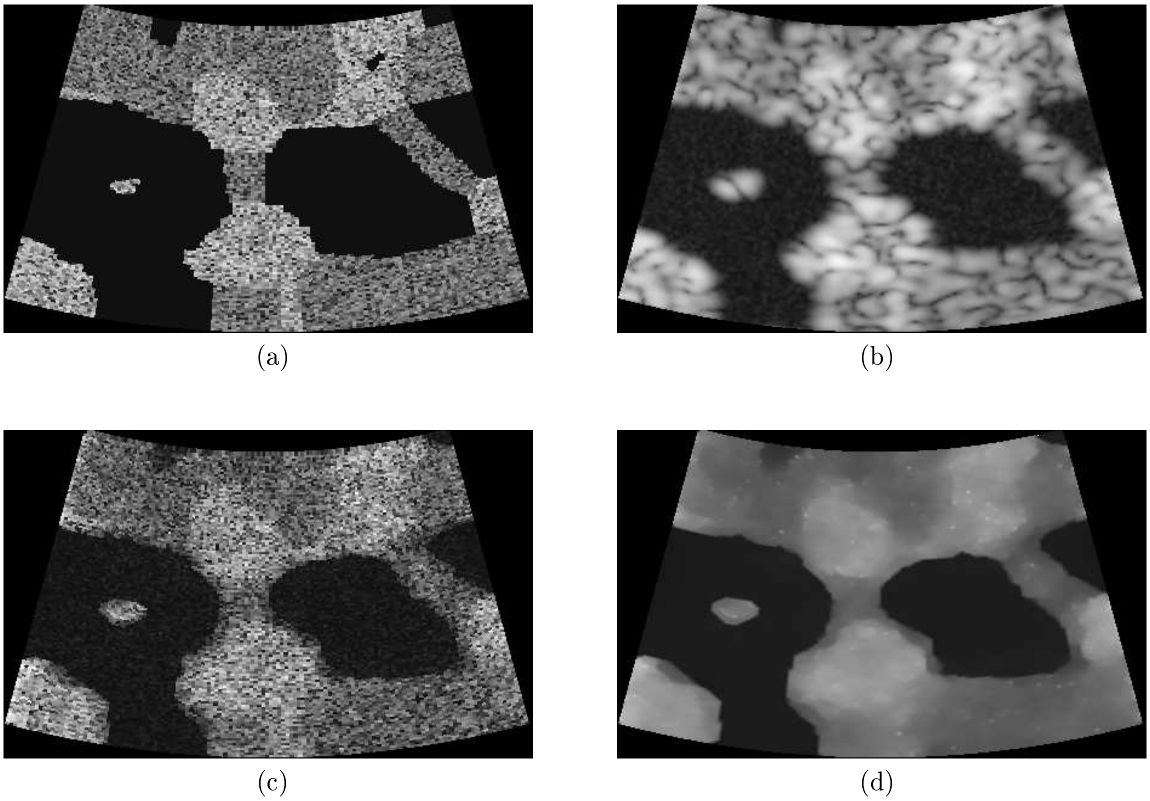

Fig. 3. Three region simulation of ultrasound image. (a) Simulated 128×128 image of diffuse scatterers r. (b) The data y produced by blurring and adding noise. (c) Estimate of reflectance field r. (d) Estimate of the variance field σ^2.

IV. Experiments

The model was tested on a simulated region image, and on a real ultrasound image; see [3] for more experiments. We implemented the algorithm in LabVIEW, and the tests were run on a Power Macintosh 8600/25 computer, using approximately 50 minutes to process each image. A ten-fold increase in speed can be expected by implementing the routines in C. Furthermore, we claim by no means optimality of the algorithm, as our emphasis was on producing best possible restorations. The processing was done in polar coordinates, but the images were log-compressed and scan-converted into Cartesian coordinates prior to display.

A. Simulated images

To produce the simulated images we make a region image [3] and use it as a variance field for the diffuse scatterers $r_{i,j}$, which are then sampled. The data y is made by convolving the ideal image with a parametric point spread function and adding white Gaussian noise. Following [18] we use a sine modulated circular Gaussian point spread function (see [3]). The radial and lateral standard deviations are estimated from the power spectrum of the observed images, while the noise variance is estimated from the power spectrum outside the acoustic bandwidth of the observed images. The parameters β, δ, ϵ, and γ are regarded as fixed, and given values $4, 2, 0.00001$ and 1, respectively. The choice of the parameters δ and γ are straightforward, guided by physical considerations (see section II-B). The parameter β was chosen by trial and error, although ideally some sort of data-driven estimation technique should have been used. See e.g. [28], [29] and [30] for approaches to parameter estimation in Markov Random Field models.

Figure 3a shows a 128×128 simulated image of diffuse scatterers with three region types having variances $0, 9$, and 27. Figure 3b shows the simulated data produced by blurring and adding white Gaussian noise with variance 0.5. The restoration in figure 3c is visually very close to the true image in 3a, although there is some smoothing due to the choice of estimator.

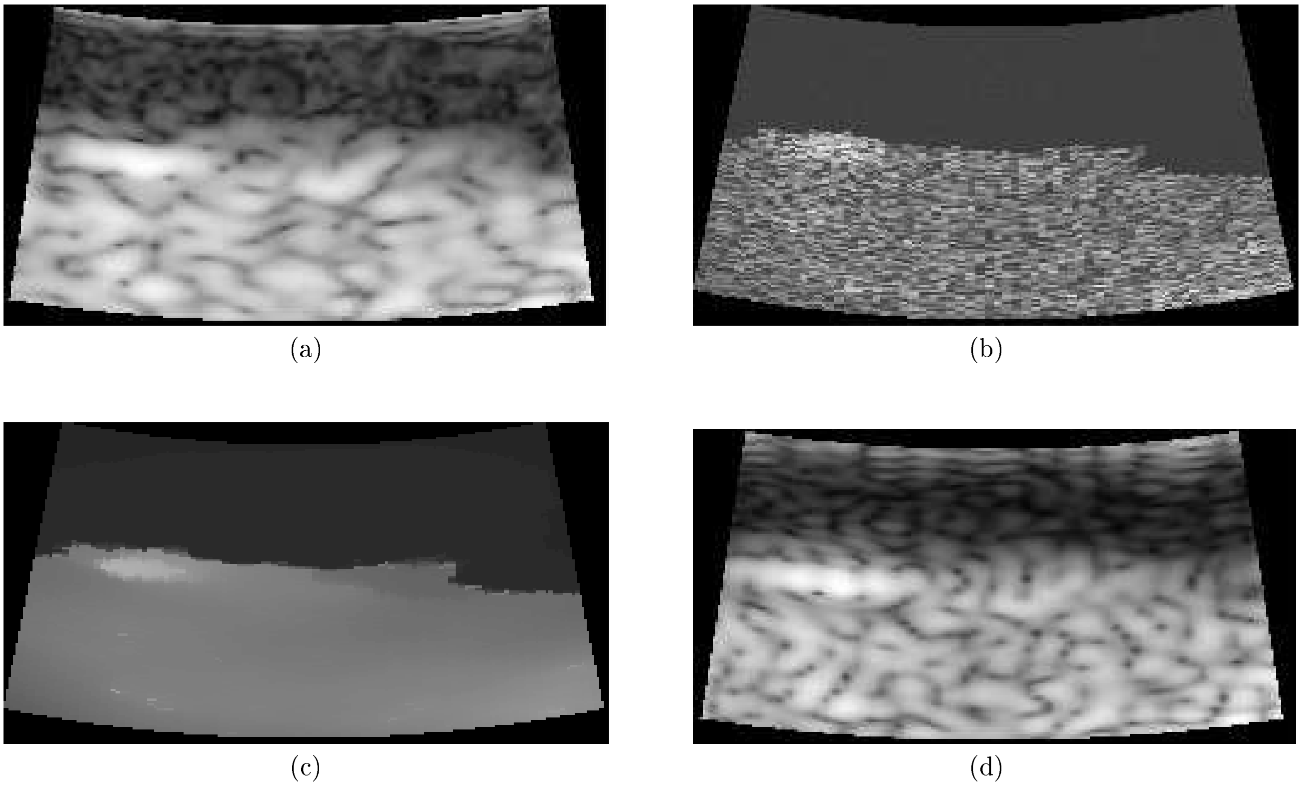

Fig. 4. Ultrasound image of a homogeneous medium submerged in water. (a) Observed 64×128 image y (b) A sample from the posterior distribution for r given y and σ^2. (c) Estimate of the variance field. (d) Wiener filtering of y.

B. Real image

The ultrasound image in figure 4 was acquired using a Vingmed Sound CFM 750 scanner with a special 16 MHz A/D-converter and a 5 MHz 15 mm annular array probe. 64 beams of 128 samples were recorded within one focal zone, and the image represent an area of approximately 1.5×0.7 cm. Figure 4a shows a homogeneous medium of diffuse scatterers submerged in water. In 4a, the lower area representing the diffuse scatterers shows much speckle, which can also be seen in the Wiener filtered image 4d. In our estimate 4b the speckle is removed, and the image shows two relatively homogeneous regions with a smooth interface. We used a sample from the posterior distribution as our estimate to avoid unnecessary smoothing. Even though the lower region of the image should be homogeneous, the point spread function has introduced artifacts that our model cannot remove. Figure 4c shows the estimated variance field, an image which in fact summarizes all relevant information in a precise fashion. The homogeneity of the two regions, and the border between them are clearly seen, without disturbing noise and image artifacts.

V. Discussion

This paper describes a model for the scattering distribution of homogeneous tissue, and introduces smoothness constraints through the use of the model in [19] for the underlying scattering characteristics. The model is used in a new algorithm for restoration of ultrasound images.

We believe that the results demonstrate the potential usefulness of Markov random field models in restoration of medical ultrasound images. The restoration of the simulated images produced results that are more representative of the true object than the observed ultrasound images. Also, in the case of the real image, the results obtained are significantly better than the one produced by Wiener filtering. In both cases the speckle patterns are efficiently removed, while at the same keeping important details and not introducing artificial structures.

Our prior model favors homogeneous areas with sharp transitions between regions, which is a reasonable characteristic for a model representing ultrasound images since tissues are mostly homogeneous with sharp transitions between different tissue types. This characteristic of the prior distribution, which can be understood by studying the plot in figure 1, is a great advantage to the restoration procedure. A disadvantage is that we may loose smaller structures. However, these structures are to a certain degree, depending on their size and the size of the degrading point spread function size, lost anyway in the blurring process.

The removal of speckle is efficient with our method, and we believe this results in visually better images. Speckle removal is important in 3D visualization systems, where the data in a 3D volume is mapped to a 2D image through a process called rendering. In order to make the 2D image understandable, it is important that only significant structure information participates in the rendering process, and that irrelevant information (i.e. noise and speckle) is excluded. It is important to notice that if a certain speckle pattern is interpreted (by the model/algorithms) as a real structure when it in fact is a homogeneous area, this can lead to misinterpretation by the end observer. This is a crucial point if the clinician wishes to use this restoration method to study and analyze details from a volume scan. In other words, if this method proves useful for real ultrasound data, a surgeon, e.g. during removal of a tumor in the brain, could perform a 3D scan and then use the method to restore the reflection image of a smaller part of the volume to investigate the remains of the tumor. It would then be crucial that the method does not produce false structures due to the speckle pattern.

Further problems have to be solved. It is well known that point spread functions in ultrasound imaging do not have Gaussian shapes, and that they vary both radially and laterally [5], [31]. As we show in [3], poor estimates of the point spread function effects the quality of the restorations. Thus effort should be put into the estimation of the point spread function in order to get truly good restorations, especially in larger images. Also, the efficiency of the algorithm should be improved. Most medical applications require real-time processing or at least on-line response times, and thus further optimization of the algorithm should be investigated. Note that when used on a sequence of images, the proceeding estimates could be used as initialization for the current frame, hopefully leading to a considerable speedup. Also, when doing real-time display of sequences, samples from the posterior distribution could be used instead of the sample means (15).

Acknowledgments

This work was partly supported by grants from the Research Council of Norway

Appendix

I. Alternative model formulation

We derive the posterior distribution $p(\sigma^2 \,|\, y, h)$, showing that our model can indeed be formulated in terms of σ^2 only. Using (3) we need the likelihood, which, using (1) and (2), is found by integrating out r:

$$p(y \,|\, \sigma^2) = \int p(y \,|\, r) p(r \,|\, \sigma^2) \mathrm{d}r. \tag{17}$$

For clarity of exposition we define matrices H and Λ as follows. Λ is a $n \times n$ diagonal matrix with entries $\sigma^2_{1,1}, \sigma^2_{1,2}, \ldots, \sigma^2_{n_r,n_l}$ on the diagonal, and H is a $n \times n$ band matrix representing the point spread function h. Letting $\Sigma_s = \tau^2 I$ we get

$$\begin{aligned}
p(y \,|\, \sigma^2) &= \int p(y \,|\, r) p(r \,|\, \sigma^2) \mathrm{d}r \\
&\propto \int \exp\left\{ -\frac{1}{2} \left((y - Hr)^T \Sigma_s^{-1} (y - Hr) - r^T \Sigma^{-1} r \right) \right\} \mathrm{d}r \\
&= \exp\left\{ -\frac{1}{2} y^T \Sigma_s^{-1} y \right\} \int \exp\left\{ -r^T \left(H^T \Sigma_s^{-1} H + \Sigma^{-1} \right) r \right. \\
&\qquad \left. + 2 r^T H^T \Sigma_s^{-1} y \right\} \mathrm{d}r \\
&= \exp\left\{ -\frac{1}{2} y^T \Sigma_s^{-1} y \right. \\
&\qquad \left. + \frac{1}{2} y^T \Sigma_s^{-1} H \left(H^T \Sigma_s^{-1} H + \Sigma^{-1} \right)^{-1} H^T \Sigma_s^{-1} y \right\} \\
&= \exp\left\{ -\frac{1}{2} y^T \left(\Sigma_s^{-1} - \Sigma_s^{-1} H \left(H^T \Sigma_s^{-1} H + \Sigma^{-1} \right)^{-1} H^T \Sigma_s^{-1} \right) y \right\} \\
&= \exp\left\{ -\frac{1}{2} y^T \left(\Sigma_s + H \Sigma H^T \right)^{-1} y \right\},
\end{aligned}$$

where the last line comes from the identity

$$\begin{aligned}
&A^{-1} - A^{-1} B \left(C^{-1} + D A^{-1} B \right)^{-1} D A^{-1} \\
&= (A + BCD)^{-1};
\end{aligned}$$

see e.g. [32] p. 459.

The posterior density is then found by combining the prior density (5) with the density of the likelihood model (7):

$$\begin{aligned}
p(\sigma^2 \,|\, y) &\propto p(y \,|\, \sigma^2) p(\sigma^2) \\
&\propto \frac{1}{\left| H \Sigma H^T + \tau^2 I \right|^{1/2}} \exp\left\{ -\frac{1}{2} y^T \left(H \Sigma H^T \;\; + \tau^2 I \right)^{-1} y - \; \Phi(\sigma^2) - \Psi(\sigma^2) \right\}.
\end{aligned} \tag{18}$$

However, note that the matrix $Q = \left(H \Sigma H^T \right)^{-1}$ in general is full; this can be seen heuristically by expanding Q as a power series. Letting $q_{ij,kl}$ be the elements of Q, it is a well known fact [33], [34] that

$$\mathrm{E}\left(Y_{i,j} \,|\, y_{-(i,j)}, \sigma^2 \right) = - \sum_{k,l \neq i,j} \frac{q_{ij,kl}}{q_{ij,ij}} \, y_{k,l}, \tag{19}$$

implying that the neighborhood of site (i, j) consists of all sites (k, l) such that $q_{ij,kl} \neq 0$. Hence since Q is full the local Markov property fails to hold since the neighborhood of each site is simply the whole image. This means that conditional on the data y, any two sites in the variance field σ^2 are directly dependent, which is undesirable for at least two reasons. First, updating is time-consuming because for each site that is updated, the full density has to be computed in (16). Secondly, because of the global dependency, the chain will move slowly in the sample space, resulting in slow convergence of the Markov chain. Thus this formulation of the model, while valid and instructive, leads to computational problems. However, this deficiency can be amended by considering the augmented field (σ^2, r), as shown in section II-C.

References

[1] T. Taxt, "Restoration of medical ultrasound images using 2-dimensional homomorphic deconvolution," *IEEE Trans. on Ultrason. Ferroelec. Frec. Contr.*, vol. 42, no. 4, pp. 543–554, June 1995.

[2] D. Iraca, L. Landini, and L. Verrazzani, "Power spectrum equalization for ultrasonic image restoration," *IEEE Trans. on Ultrason. Ferroelec. Frec. Contr.*, vol. 36, no. 2, pp. 216–222, march 1989.

[3] T. Langø, T. Lie, O. Husby, and J. Hokland, "Bayesian 2D deconvolution: Effect of non-stationary ultrasound pulses," *Submitted to IEEE Trans. on Ultrason. Ferroelec. Frec. Contr*, 1999.

[4] T. Taxt and G. Frolova, "Noise robust one-dimensional blind deconvolution of medical ultrasound images," *IEEE Trans. on Ultrason. Ferroelec. Frec. Contr.*, vol. 46, no. 2, pp. 291–299, march 1999.

[5] J.A. Jensen and S. Leeman, "Nonparametric estimation of ultrasound pulses," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 10, pp. 929–936, 1994.

[6] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.

[7] X. Zong, A.F. Laine, and E.A. Geiser, "Speckle reduction and contrast enhencement of echocardiograms via multiscale nonlinear processing," *IEEE Transaction on Medical Imaging*, vol. 17, no. 4, pp. 532–540, August 1998.

[8] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Springer, Berlin, 1995.

[9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.

[10] J. Besag, P. Green, D. Higdon, and K. Mengersen, "Bayesian computation and stochastic systems (with discussion)," *Statistical Science*, vol. 10, no. 1, pp. 3–66, 1995.

[11] J.W. Goodman, "Statistical properties of laser speckle patterns," in *Laser Speckle and Related Phenomena*, J.C. Dainty, Ed. Springer Verlag, Berlin, 1975.

[12] R.F. Wagner, M.F. Insana, and D.G. Brown, "Statistical properties of radio-frequency and envelope-detected signals with application to medical ultrasound," *J. Opt. Soc. Am, A*, vol. 4, no. 5, pp. 910–922, 1987.

[13] S. Geman and D.E. McClure, "Statistical methods for tomographic image reconstruction," in *Proc. 46th Sess. Int. Stat. Inst. Bulletin ISI*, 1987, vol. 52.

[14] W. K. Hastings, "Monte Carlo simulation methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[15] P. J. Green, "Reversible jump MCMC computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.

[16] J. Besag and P. Green, "Spatial statistics and Bayesian computation," *Journal of the Royal Statistical Society, Series B*, vol. 16, pp. 395–407, 1993.

[17] D.M. Higdon, "Auxiliary variable methods for Markov chain Monte Carlo with applications," *Journal of the American Statistical Association*, vol. 93, no. 442, pp. 585–595, June 1998.

[18] J.H. Hokland and P.A. Kelly, "Markov models of specular and diffuse scattering in restoration of medical ultrasound images," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 43, no. 4, pp. 660–669, July 1996.

[19] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 3, pp. 367–383, 3 1992.

[20] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Transaction on Medical Imaging*, vol. 6, no. 2, pp. 298–311, february 1997.

[21] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman & Hall, 1996.

[22] M. Hurn and C. Jennison, "An extension of Geman and Reynolds' approach to constrained restoration and the recovery of discontinuities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 6, pp. 657–662, June 1996.

[23] D. Geman, G. Reynolds, and C. Yang, "Stochastic algorithms for restricted image spaces and experiments in deblurring," in *Markov Random Fields: Theory and Application*, R. Chellappa and A. Jain, Eds. Academic Press, New York, 1993.

[24] G.O. Roberts, A. Gelman, and W.R. Gilks, "Weak convergence and optimal scaling of random walk metropolis algorithms," *Ann. Appl. Prob.*, vol. 7, pp. 110–120, 1997.

[25] S. P. Brooks and G. O. Roberts, "Assessing convergence of Markov chain Monte Carlo algorithms," *Statistics and Computing*, vol. 8, no. 4, pp. 319–335, 1998.

[26] C. P. Robert, "Convergence control methods for MCMC algorithms," *Statistical Science*, vol. 10, no. 3, pp. 231–253, 1995.

[27] K. L. Mengersen, C.P. Robert, and C. Guihenneuc-Jouyaux, "MCMC convergence diagnosics: a 'reviewww'," Tech. Rep., CREST, INSEE, Paris, May 1998.

[28] C. Geyer and E. A. Thompson, "Constrained Monte Carlo maximum likelihood for dependent data," *Journal of the Royal Statistical Society, Series B*, vol. 54, no. 3, pp. 657–699, 1992.

[29] A. M. Thompson, J. C. Brown, . W. Kay, and D. M. Titterington, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 326–339, april 1991.

[30] D. M. Higdon, J. E. Bowsher, V. E. Johnson, T. G. Turkington, D. R. Gilland, and R. J. Jaszczak, "Fully Bayesian estimation of Gibbs hyperparameters for emission computed tomography data," *IEEE Transaction on Medical Imaging*, vol. 14, no. 5, pp. 516–526, october 1997.

[31] L.A. Ødegård, *Phase aberration correction in medical ultrasound imaging*, Ph.D. thesis, The Norwegian Institute of Technology, 1995.

[32] K.V. Mardia, J.T. Kent, and J.M. Bibby, *Multivariate Analysis*, Academic Press, 1979.

[33] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society, Series B*, vol. 36, pp. 192–236, 1974, with discussion.

[34] J. Besag and C. Kooperberg, "On conditional and intrinsic autoregressions," *Biometrika*, vol. 82, no. 4, pp. 733–746, december 1995.