Optical imaging in a variational Bayesian framework

S Arhab*, H Ayasso2, B Duchêne1 and A Mohammad-Djafari1

1 Laboratoire des Signaux et Systèmes (L2S, UMR8506: CNRS - SUPELEC - Univ Paris Sud)
2 Grenoble Images Parole Signal Automatique (GIPSA-lab, UMR5216: CNRS - Grenoble-INP, Univ Joseph Fourier - Univ Stendhal)

E-mail: slimane.arhab@lss.supelec.fr

Abstract. We are interested in optical imaging of nano-structured man-made objects. Optical imaging is taken as a nonlinear inverse scattering problem where the goal is to retrieve the dielectric parameters of an unknown object. In addition to be nonlinear, such problems are also known to be ill-posed, which means that a regularization is required prior to their resolution. This is done by introducing a priori information which consists in the fact that object is known to be composed of compact homogeneous regions made of a finite number of different materials. This a priori knowledge is appropriately translated in a Bayesian framework by a Gauss-Markov-Potts prior. Hence, a Gauss-Markov random field is used to model the contrast distribution, whereas a hidden Potts-Markov field accounts for the compactness of the regions. The problem is then solved by means of a variational Bayesian approximation, which consists in approximating the joint posterior law of all unknown parameters in the Kullback-Leibler sense with a separable free form distribution. This leads to an implicit parametric optimization scheme which is solved iteratively. This inversion algorithm is applied to laboratory controlled experimental data.

1. Introduction

Optical Digital Tomographic Microscopy (ODTM) is a quantitative imaging technique [1] in which the opto-geometrical parameters of an object (shape, permittivity) can be determined with a good resolution from far field measurements. Thanks to the synthetic high numerical aperture obtained by varying the incidence angle, the lateral resolution achieved is at least twice better than that obtained with classical far field microscopes. The opto-geometrical parameters are determined by applying numerical inversion algorithm, with an important role played by the multiple scattering regime in terms of resolution. In classical optical microscopy, the latter is limited by the Rayleigh-Abbe criterion. However, when highly refractive objects are considered, resolutions beyond the Abbe-Rayleigh criterion can be achieved by accounting for multiple scattering through a rigorous resolution of Maxwell’s equations. Optical microscopy is then taken as a non linear inverse scattering problem described through two coupled contrast source domain integral equations obtained by applying Green’s theorem to the Helmholtz wave equations satisfied by the fields and by accounting for boundary and radiation conditions.

In the early 90’s, several deterministic inversion algorithms [2, 3, 4] have been developed in order to deal with the nonlinear problem at hand, through an iterative minimization of a cost functional that accounts for both integral equations mentioned above and that expresses the discrepancy between the measured scattered fields and the fields associated to the best available solution (the sought contrast). However, in addition to be nonlinear, inverse scattering problems are also known to be ill-posed, which means that a regularization is required prior to their resolution. Such a regularization is usually done by introducing a priori information on the object, which is not easy with the latter techniques as this information must be introduced in the functional to be minimized. On the contrary, the probabilistic Bayesian framework allows us to take easily into account such information.
In the present case, we consider man-made objects that are known to be composed of compact homogeneous regions made of a finite number of different materials. This a priori is introduced through a Gauss-Markov-Potts model [5, 6], the marginal distribution of contrast being sought for as a Gaussian mixture [7] where each Gaussian law represents a class of materials, and the compactness of the regions being accounted for by a hidden Markov model.

In Bayesian methods the joint posterior distribution of the unknowns and parameters of the model is sought for by means of the Bayes rule. Generally, the estimators are computed in the maximum a posteriori (MAP) or posterior mean (PM) senses that are hardly tractable in the present case. Therefore, an approximation of the posterior law is introduced in order to obtain the desired solution at a reasonable computational cost. The latter is based upon the deterministic Variational Bayesian Approach (VBA, [9]) inspired by statistical physics [8], that approximates the posterior distribution by a separable density. As the calculations are then analytic, the rate of convergence of VBA is expected to be fast, at least much faster than that of the traditional MCMC approaches that are rather costly in terms of computation time as they attempt to estimate the posterior distribution by performing an empirical mean through a stochastic Gibbs sampling [10].

Such a method has been developed and successfully applied in the framework of microwave imaging [11]. It has also been successfully applied to optical imaging in a configuration different from that considered herein [12, 13] that leads to a different observation operator.

2. The experimental configuration and direct scattering problem

The object considered herein, whose cross-section $\Omega$ is depicted in Figure 1, is made of two parallel 25-nm-height 50-nm-width germanium rods of long extent lying on a glass substrate and 50-nm distant from one other. The substrate is of known relative permittivity and its dimensions are large as compared to those of the rods, so that the configuration is modeled as follows: an object made of germanium lies in the upper layer of a stratified medium made of two semi-infinite half-spaces separated by a planar interface $\gamma_{12}$. The upper half-space $\mathcal{D}_1$ is air and the lower one $\mathcal{D}_2$ is glass. The different media are supposed to be lossless and they are characterized by their propagation constant $k_m$ ($m = 1, 2$ or $\Omega$) such that $k_m^2 = \omega^2 \varepsilon_0 \varepsilon_m \mu_0$, where $\omega$ is the angular frequency, $\varepsilon_0$ ($\varepsilon_0 = 8.854 \times 10^{-12}$ F m$^{-1}$) and $\mu_0$ ($\mu_0 = 4\pi \times 10^{-7}$ H m$^{-1}$) are the dielectric permittivity and the magnetic permeability of free space, respectively, and $\varepsilon_m$ is the relative dielectric permittivity of medium $\mathcal{D}_m$ ($\varepsilon_1 = 1$, $\varepsilon_2 = 2.25$ and $\varepsilon_\Omega = 29.8$). It can be noted that $\varepsilon_\Omega$ and $\Omega$ are supposed to be unknown in the inversion process.

The object is supposed to be contained in a test domain $\mathcal{D}$ ($\mathcal{D} \subset \mathcal{D}_1$) and we introduce a contrast function $\chi$ such that $\chi(r) = (k^2(r) - k_1^2)/k_1^2$, defined in $\mathcal{D}$ and null outside $\Omega$. The object is illuminated by an incident wave generated by a Helium-Neon laser, operating at a 633 nm wavelength, coupled with a reflection microscope equipped with an interferometric device able
to provide accurate measurements of the phase of the scattered fields [14]. The incident wave, with an implicit time-dependence \( \exp(-i \omega t) \), can be considered as a plane wave whose electric field \( E^{\text{inc}} \) is polarized in the \( \gamma_{12} \) interface plane along the axis of the rods so that a 2D-TM configuration is considered. The incident wave illuminates the rods from the glass side in direction \( \theta_1 \) that can be varied in the range \( \pm 50^\circ \); hence \( N_l = 9 \) views are carried out at varying \( \theta_1 \), each view being constituted of measurements of the scattered field in the far field domain \( S \) at \( N_r = 683 \) different observation angles \( \theta \) in the range \( \pm 54^\circ \).

Modeling of the wave object interaction is done by means of two coupled domain integral equations obtained by applying Green’s theorem to the Helmholtz wave equations satisfied by the fields and by accounting for boundary and radiation conditions. The first one, denoted as observation equation, relates the scattered electric field \( y(r') \) observed on a measurement domain \( S (r' \in S) \) to the Huygens type sources \( w(r) \) induced within the object \( (r \in D) \) by the incident wave, i.e. \( w(r) = \chi(r)E(r) \) where \( E(r) \) is the total field in the object, and involves Green’s function of the stratified medium with source in \( D_1 \) and observation in \( D_2 \). The second one, the so-called coupling (or state) equation, relates the total field in the object to the induced sources \( w(r) \) and the incident field \( E^{\text{inc}}(r) \) and involves Green’s function with both source and observation located in \( D_1 \) [15, 16]. The forward problem then consists in first solving the coupling equation for \( E \), knowing \( \chi \), then solving the observation equation for \( y \) knowing \( w \). This is done from discrete counterparts of these equations obtained by means of a method of moments with pulse-basis and point matching. The test domain \( D \) is then partitioned into \( N_D \) elementary square pixels with side \( \delta \) small enough to consider the field and the contrast as constant over each of them. Figure 2 displays the results obtained in this way compared to experimental data collected at Institut Fresnel in a controlled situation and a good agreement can be observed between both fields, once omitted the high level of noise in the experimental data that can be explained by the fact that the amplitude of the scattered field is very low as compared to that of the incident field.

![Figure 2](image-url)

**Figure 2.** Modulus (a) and phase (b) of the calculated (dashed line) and measured (full line) scattered fields for the illumination \( \theta_i = 49^\circ \).

### 3. The inverse scattering problem and results

Let us now introduce two variables, \( \epsilon \) and \( \xi \) that account for uncertainties and measurement and model errors and that are supposed to be centred and white and to satisfy Gaussian laws. The observation and coupling equations then read in a condensed form:

\[
\begin{align*}
y_l &= \mathcal{H}^S w_l + \epsilon_l \\
w_l &= \chi E^{\text{inc}}_l + \chi \mathcal{H}^D w_l + \xi_l,
\end{align*}
\]
where $\mathcal{H}^S$ and $\mathcal{H}^D$ are matrices whose elements result from the integration over the elementary pixels of the observation and coupling Green’s functions, respectively, and subscript $l$ accounts for the different views.

Let us now introduce a priori information that the object is composed of a finite number $N_z$ of different materials. This is done by means of a hidden variable $z_r$ ($z_r = z(r)$, $r = 1, \ldots, N_r$) that associates a class of materials to each pixel $r$. Hence, when $z(r) = k$, the contrast value of pixel $r$ is given by a Gaussian distribution $\mathcal{N}(\mu_k, \nu_k)$ with mean value $\mu_k$ and variance $\nu_k$, the statistical parameters of the $k^{th}$ class. The object contrast $\chi$ is then defined by a multivariate Gaussian distribution that follows a conditionally independent Gauss mixture model:

$$p(\chi|z, \mu, \nu) = \prod_{k=1}^{N_z} \frac{1}{(2\pi \nu_k)^{N_k/2}} \exp\left\{ -\frac{1}{2\nu_k} \sum_{r \in R_k} ||\chi_r - \mu_k||^2 \right\},$$

where $N_k = \text{card}(R_k)$ is the number of pixels that belong to the $k^{th}$ class. It can be noted that pixels are conditionally independent between each other.

A priori information that the different materials are distributed in compact homogeneous regions is accounted for by means of a Potts-Markov model on $z$:

$$p(z|\gamma) = \frac{1}{\Omega} \exp\left\{ \sum_{r=1}^{N_r} \sum_{r' \in O(r)} \delta[z(r) - z(r')] \right\},$$

where $O(r)$ is a neighborhood of $r$, herein made of its four nearest neighbors, $\Omega$ is a normalization constant and $\gamma$ determines the correlation between neighbors.

The conditional probability laws for scattered fields and induced sources are Gaussian distributions that account for the observation and coupling equations, respectively. They read:

$$p(y|w, \nu_{\epsilon}) = \frac{1}{(2\pi \nu_{\epsilon})^{(N_r N_l)/2}} \exp\left\{ -\frac{1}{2\nu_{\epsilon}} \sum_{l=1}^{N_l} ||y_l - \mathcal{H}^S w_l||^2_S \right\}$$

$$p(w|x, \nu_{\xi}) = \frac{1}{(2\pi \nu_{\xi})^{(N_r N_l)/2}} \exp\left\{ -\frac{1}{2\nu_{\xi}} \sum_{l=1}^{N_l} ||w_l - \chi E^\text{inc} - \chi \mathcal{H}^D w_l||^2_D \right\},$$

where $||.||_A$ is the norm associated to the inner product in $L_2(A)$, $A = S$ or $D$.

From now on, all the hyper-parameters are grouped in the set $\theta = \{\nu_{\epsilon}, \nu_{\xi}, \mu, \nu\}$ and conjugate probability laws are assigned to them. Hence $p(\theta) = p(\nu_{\epsilon}) p(\nu_{\xi}) p(\mu) p(\nu)$, with $p(\mu) = \prod_k p(\mu_k)$ and $p(\nu) = \prod_k p(\nu_k)$, where the probability laws for the mean values $\mu_k$ are Gaussian distributions whereas that of the variances $\nu_k$ are inverse-gamma distributions as well as that of $\nu_{\epsilon}$ and $\nu_{\xi}$:

$$p(\mu_k|\mu_o, \nu_o) = \frac{1}{(2\pi \nu_o)^{1/2}} \exp\left\{ -\frac{1}{2\nu_o} |\mu_k - \mu_o|^2 \right\} \quad k = 1, \ldots, N_z$$

$$p(\nu_k|\alpha_o, \beta_o) = IG(\alpha_o, \beta_o) = (1/\Gamma(\alpha_o)) \beta_o^{\alpha_o} \nu_k^{-(\alpha_o+1)} \exp(-\beta_o/\nu_k) \quad k = 1, \ldots, N_z$$

$$p(\nu_{\epsilon}|\alpha_{\epsilon}, \beta_{\epsilon}) = IG(\alpha_{\epsilon}, \beta_{\epsilon}) = (1/\Gamma(\alpha_{\epsilon})) \beta_{\epsilon}^{\alpha_{\epsilon}} \nu_{\epsilon}^{-(\alpha_{\epsilon}+1)} \exp(-\beta_{\epsilon}/\nu_{\epsilon}) \quad \epsilon = \epsilon, \xi.$$  

The meta-hyper-parameters $\mu_o, \nu_o, \alpha_o, \beta_o, \alpha_{\epsilon}$ and $\beta_{\epsilon}$ ($\tau = \epsilon, \xi$) are set to satisfy non-informative flat prior distributions.

It can be noted that an unsupervised context is considered here and the contrast, contrast sources and hyper-parameters of the model are estimated at the same time. The joint posterior distribution is then obtained by applying the Bayes rule:
p(w, χ, z, θ|y) ∝ p(y|w, ν) p(w|χ, ν_χ) p(χ|z, μ, ν) p(z|γ) p(θ).

From this joint posterior law, different inferences can be done. Usually an estimator such as the maximum a posteriori (MAP) or the posterior mean (PM) is sought for. In the present case, these estimators are very difficult to obtain. This is the reason why a variational Bayesian method is adopted. The latter has been developed in order to perform approximate posterior inferences at a low computational cost compared to stochastic sampling methods [17, 18]. It consists in approximating the joint posterior law by a free form separable distribution where the unknowns are updated iteratively so as to minimize the Kullback-Leibler divergence:

$$\text{KL}(\tilde{p}(\Theta|f)||p(\Theta|f)) = \int \tilde{p}(\Theta|f) \log[\tilde{p}(\Theta|f)/p(\Theta|f)] \, d\Theta,$$

where $$\Theta = \{w, \chi, z, \theta\}$$. The free form distribution then reads [9]:

$$\tilde{p}(w, \chi, z, \theta|y) = \prod_{l=1}^{N_l} \prod_{r=1}^{N_\chi} \tilde{p}(w_l(r)|y) \prod_{r=1}^{N_\chi} \tilde{p}(\chi(r)|y) \prod_{r=1}^{N_z} \tilde{p}(z(r)|y) \prod_{k=1}^{N_\mu} \tilde{p}(\mu_k|y) \times \prod_{k=1}^{N_\nu} \tilde{p}(\nu_k|y) \, \tilde{p}(\nu_\chi|y) \, \tilde{p}(\nu_\xi|y),$$

and is detailed in [13].

Figure 3 displays the results obtained by means of the above method, i.e. the behavior of the means and variances of the contrast of classes 1 and 2 during the iterative process and the contrast obtained after 600 iteration steps. In spite of the high level of noise that appears in the scattered field data, the algorithm succeeds in retrieving the two germanium rods, but with a better reconstruction for the left rod than for the right one, which can be explained by an unexpected dissymmetry in the noise level as it appears in figure 2.

**Figure 3.** Behavior of the mean $$m_k$$ and variance $$\nu_k$$ of the contrast for classes $$k=1$$ (air) and $$k=2$$ (germanium) during iterations (left) and the contrast retrieved by means of the variational Bayesian approach (right).

4. Conclusion
The results presented in this paper are a step toward the development of a new optical tomographic microscopy technique that yields a nanometric resolution far beyond that of standard optical microscopes limited by the Abbe-Rayleigh criterion. The dimensions of the object studied herein have been chosen in order to show that. Of course objects with smaller details could certainly be imaged. However, resolution limit achievable with this technique is
not easy to quantify as it depends upon the level of noise that corrupts the data and upon the object itself. Indeed, the imaging technique is based upon a nonlinear inversion algorithm that benefits from multiple scattering due to the high dielectric contrast of the sought object and such a good resolution could certainly not be obtained with a lowly contrasted object imaged by means of a standard Born-based inversion method. On the other hand, deterministic nonlinear inversion methods such as the Newton-Kantorovich algorithm [2], the modified gradient method [3] or the contrast source inversion [4], are known to be limited in contrast level that can be accurately quantitatively retrieved, essentially due to the fact that they are based upon local minimization schemes. The method presented herein does not show such a drawback as it is not based upon such a strategy and as a priori information on the sought object introduced in the inversion algorithm allows us to reduce the eligible solution space. This can be quantified and constitutes the next step of this work, namely "what does a priori information add in terms of image quality?", as, from now on, it is more appropriate to quantify this quality rather than the standard resolution. Another interesting feature remains to be studied, i.e. the inversion of multifrequency data, that would certainly allows us to get a better image quality.

5. Acknowledgments
This work has been supported by Agence National de la Recherche under the convention n° ANR 12 BS03 006 02 (project “SURMITO”). The authors would like to acknowledge G. Maire, K. Belkebir and A. Santenac (Institut Fresnel, Marseille) for the experimental data.

6. References
[1] Lauer V 2002 New approach to optical diffraction tomography yielding a vector equation of diffraction tomography and a novel tomographic microscope J. Microsc. 205(2) pp 165–176
[2] Joachimowicz , Pichot C and Hugonin JP 1991 Inverse scattering: an iterative numerical method for electromagnetic imaging IEEE Trans. Antennas Propag. 39(12) pp 1742–1752
[3] Kleinman RE and van den Berg PM 1992 A modified gradient method for two-dimensional problems in tomography J. Comput. Appl. Math. 42(1) pp 17–35
[4] van den Berg PM and Kleinman RE 1997 A contrast source inversion method Inverse Prob. 13(6) pp 1607–1620
[5] Tierney L 1994 Markov chains for exploring posterior distributions The Annals of Statistics 22(4) pp 1701–1728
[6] Pieczynski W 2003 Markov models in image processing Traitement du Signal 20(3) pp 255–278
[7] Férón O and Mohammad-Djafari A 2002 Image fusion and joint segmentation using an MCMC algorithm J. Electron. Imaging 14 023014
[8] MacKay DJC 1995 Ensemble learning and evidence maximization Proc. NIPS 10.1.1.54.4083
[9] Smidl V and Quinn A 2005 The Variational Bayes Method in Signal Processing (New York: Springer)
[10] Robert C and Casella G 2005 Monte Carlo Statistical Methods (New York: Springer)
[11] Férón O, Duchêne B and Mohammad-Djafari A 2005 Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data Inverse Prob. 21(6) pp 895–9115
[12] Ayasso H, Duchêne B and Mohammad-Djafari A 2010 Bayesian inversion for optical diffraction tomography J. Mod. Opt. 57(9) pp 765–776
[13] Ayasso H, Duchêne B and Mohammad-Djafari A 2012 Optical diffraction tomography within a variational Bayesian framework Inverse Probl. Sci. Eng. 20(1) pp 59–73
[14] Girard J, Maire G, Giovannini H, Talneau A, Belkebir K, Chaumet PC and Sentenac A 2010 Nanometric resolution using far-field optical tomographic microscopy in the multiple scattering regime Phys. Rev. A 82(6), 061801
[15] Lesselier D and Duchêne B 1991 Buried, 2-D penetrable objects illuminated by line sources: FFT-based iterative computations of the anomalous field Application of Conjugate Gradient Methods to Electromagnetics and Signal Analysis (Progress In Electromagnetics Research vol 5) ed TK Sarkar (New York: Elsevier) pp 351–389
[16] Grefet JJ 1989 Scattering of s-polarized electromagnetic waves by a 2D obstacle near an interface Opt. Commun. 72(5) pp 274–278
[17] Beal MJ 2003 Variational algorithms for approximate Bayesian inference University of London PhD thesis
[18] Jaakkola TS 2001 Tutorial on variational approximation methods Advanced Mean Field Methods: Theory and Practice eds M Opper and D Saad (Cambridge: MIT Press) pp 129–159