BAYESIAN REGULARIZATION IN NON-LINEAR IMAGING: RECONSTRUCTIONS FROM EXPERIMENTAL DATA IN MICROWAVE TOMOGRAPHY

Roberta Autieri (1), *Giancarlo Ferraiuolo* (2), *and Vito Pascazio* (1)

(1) Dipartimento per le Tecnologie, Università di Napoli "*Parthenope*"
via Medina 40, 80133, Napoli, Italia
phone: + 39 081 5474889, fax: + 39 081 5474777, email: {roberta.autieri,vito.pascazio}@uniparthenope.it
web: www. ingegneria.uniparthenope.it

(2) Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni, Università di Napoli "*Federico II*"
via Claudio 21, 80125, Napoli, Italia
phone: + 39 081 7683810, fax: + 39 081 7683149, email: gferraiu@unina.it
web: www.diet.unina.it

ABSTRACT

In this paper we investigate the robustness and the effectiveness of a microwave imaging technique, based on Bayesian estimation theory, for the reconstruction of dielectric profiles. The validation is conducted on real experimental data, the well-known "Marseille" dataset. Our statistical based inversion algorithm takes advantage of Bayesian regularization, which permits to invert a strongly non-linear model using a Markov Random Field as a-priori statistical model of the unknown image. Such choice leads to a robust and effective non-linear inversion method. An exhaustive analysis on the experimental data is also performed, in order to show the good performance of the method.

1. INTRODUCTION

Tomographic imaging at microwaves amounts to produce images (*tomograms*) of internal sections of objects [1]. Such images are usually formed by the samples of the so-called contrast function, that is strictly related to the complex permittivity of the object. Tomographic imaging can be useful in wide range of applications, for instance in medical imaging of human body, in the detection of internal defeats of objects used in aircrafts and nuclear plants, in ground penetrating radar (GPR) imaging for archeology, for underground tunnels detection, and many others.

Microwave tomography is a typical case of inverse problem for image formation. The images are obtained by illuminating the objects with known incident fields and by measuring the fields they scatter outside. The measured scattered fields data samples are related to the contrast function (the unknown) through a non linear mapping. Since the overall information content of the measured scattered fields data is bounded, and the unknown samples belong to functional spaces whose dimension is not finite, the tomographic problem is ill-posed and a regularization strategy is necessary [2]. Typically, this consists of imposing some constraints to the solution, and it is usually realized by using a penalty term.

Basically, regularization techniques can be grouped in two main categories: in the first one the unknowns are modelled as a deterministic function (classic approach), in the second the unknowns are modelled as a random process (Bayesian approach). Classic techniques make use of quadratic (Tikhonov), entropy-type, or roughness measure penalty terms. Regardless of this choice, a crucial point is the evaluation of the tuning parameter related to the regularizing term; in a non-linear framework this task becomes very difficult and generally is overcame through supervised and empirical techniques. Bayesian techniques make use of statistical a-priori information. A question common to all approaches is related to the need of using iterative techniques to minimize the adopted cost function. Being the mapping between data and unknowns non-linear, such cost function is non-convex. In this situation, the convergence of the algorithm depends on the starting point of the procedure, and the eventuality of trapping in a local minima, representative of a false solution, is not negligible.

In this paper, we present a Markov Random Field (MRF) based Maximum A Posteriori (MAP) estimation technique (which belongs to the class of Bayesian approaches) to solve this problem, and we show how the value of the regularization parameters can be estimated avoiding supervised and empirically based solutions.

The paper is organized as follows: in the next section we present the microwave tomography model, and the inversion procedure, consisting of both image and parameter's estimation algorithms. In section 3 we present a performance analysis based on the inversion of a well known experimental data set [3]. In section 4, conclusions and future works are outlined.

2. THE MICROWAVE TOMOGRAPHY MODEL AND THE INVERSION PROCEDURE

To recover faithful tomograms it is necessary to invert a mathematical model relating known quantities (the probing incident field E_i, and the measured scattered fields E_s), to the unknown ones (the object internal characteristics we are interested to, such as the dielectric permittivity ε, and the total electric field E inside the object). Such a model in a two dimensional (2-D) geometry is given by [4]:

$$E_s^v(\boldsymbol{R}_m)=\int_\Omega g_{\text{ext}}(\boldsymbol{r}-\boldsymbol{R}_m)\chi(\boldsymbol{r})E^v(\boldsymbol{r})d\boldsymbol{r}+n(\boldsymbol{R}_m),\quad \boldsymbol{R}_m\notin\Omega,$$
$$E^v(\boldsymbol{r})=E_i^v(\boldsymbol{r})+\int_\Omega g_{\text{int}}(\boldsymbol{r}'-\boldsymbol{r})\chi(\boldsymbol{r}')E^v(\boldsymbol{r}')d\boldsymbol{r}',\qquad \boldsymbol{r}\in\Omega, \tag{1}$$

where Ω is the compact support of the object, $\boldsymbol{R}_m$, m=1,…,M, is the co-ordinate of the m-th measurement point, $\chi(.)=[\varepsilon_r(.)-1]$ ($\varepsilon_r(.)$ is the complex relative dielectric permittivity of the object) is the so-called "contrast function", the *unknown* tomographic image, $g_{\text{ext}}(.)$ and $g_{\text{int}}(.)$ are the Green's functions to be chosen according to the geometry of the problem [4], and the additive term $n(.)$ is a zero mean white Gaussian noise (AWGN) process with known variance σ_n^2. The index v, where v=1,…,V, denotes the V different illumination conditions: the probing incident fields can be changed, so that new fresh data from different aspect angles can be collected [4].Consequently, a total of MxV measurement scattered field values are collected.

Equations (1) cannot be solved analytically, and must be properly discretized as shown in Ref [4]. After discretization of eqs. (1):

$$\boldsymbol{y}=\mathbf{A}_e(\boldsymbol{x})\boldsymbol{e}+\boldsymbol{n}$$
$$\boldsymbol{e}_i=[\mathbf{I}-\mathbf{A}_i(\boldsymbol{x})]\boldsymbol{e} \tag{2}$$

where $\mathbf{A}_e$ and $\mathbf{A}_i$ are the discrete counterparts of integrals in eqs. (1), and $\boldsymbol{y}$, $\boldsymbol{e}_i$, $\boldsymbol{e}$, $\boldsymbol{x}$ and $\boldsymbol{n}$ are vectors collecting the samples of scattered field, incident field, total field, contrast, and noise, respectively, and after substitution of the second equation into the first, the tomographic problem can be synthetically formulated as [5]:

$$\boldsymbol{y}=\mathbf{A}_e(\boldsymbol{x})[\mathbf{I}-\mathbf{A}_i(\boldsymbol{x})]^{-1}\boldsymbol{e}_i+\boldsymbol{n}=\mathbf{A}(\boldsymbol{x})+\boldsymbol{n}. \tag{3}$$

The general formulation of the problem solution in the classical statistical framework leads to the Maximum Likelihood (ML) estimation method. In particular, in our case (AWG noise), ML reduces to least square (LS) estimation. However, because of the ill-posedness of the inverse problem, the solution is typically achieved minimizing or maximizing a cost function, respect to the unknowns [2]:

$$\psi(\boldsymbol{y},\boldsymbol{x},\boldsymbol{x}_0)=\phi(\boldsymbol{x},\boldsymbol{y})+\alpha\phi_d(\boldsymbol{x},\boldsymbol{x}_0) \tag{4}$$

where $\phi(\boldsymbol{x},\boldsymbol{y})$ is a term related to the likelihood function (in our case $\phi(\boldsymbol{x},\boldsymbol{y})=(1/2\sigma_n^2)\|\boldsymbol{y}-\boldsymbol{A}(\boldsymbol{x})\|^2$, which is also the function to be minimized in the LS problem), $\phi_d(\boldsymbol{x},\boldsymbol{x}_0)$ is the penalty function, that is a measure of the distance between the estimate and a nominal value $\boldsymbol{x}_0$, and α is the Lagrange multiplier.

MAP estimation technique requires the use of an a-priori statistical model for the image to be estimated. We adopt a Gaussian MRF (GMRF) as a-priori model [6], in a complex form adapted on the current complex formulation of the problem. In this model, we use two different regularization parameters, from here called "hyperparameters", one for the real part of the image and another for the imaginary part. The hyperparameters are estimated from the measured data. The method requires the use of the Expectation-Maximization (EM) algorithm and the Metropolis algorithm [7], optimised for the case [8]. After hyperparameter's estimation, it is possible to perform the inversion via MAP estimation. This is obtained by a sub-optimal deterministic minimization technique to achieve high resolution capabilities at low computational cost.

2.1 Image Estimation Procedure

Following the Maximum a Posteriori (MAP) formulation, and denoting with $f_{\boldsymbol{X}|\boldsymbol{Y}}$ the *a posteriori* probability density function (*pdf*) of the unknown, and with $f_{\boldsymbol{X}}$ the *a priori* pdf, the inversion procedure consists of finding the solution that satisfies the following criterion:

$$\hat{\boldsymbol{x}}_{MAP}=\arg\max_{\boldsymbol{x}} f_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y})=\arg\max_{\boldsymbol{x}} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x}), \tag{5}$$

where, in this case, both data and unknowns are considered as samples of a random process, denoted by $\boldsymbol{X}$ and $\boldsymbol{Y}$. With reference to model (3), and in case of AWG noise, we get:

$$\hat{\boldsymbol{x}}_{MAP}=\arg\min_{\boldsymbol{x}}\left[\frac{1}{2\sigma_n^2}\|\mathbf{A}(\boldsymbol{x})-\boldsymbol{y}\|^2-\log f_{\boldsymbol{X}}(\boldsymbol{x})\right], \tag{6}$$

Implementation of (6) requires the knowledge of the a priori pdf f_X. A common choice in the image processing community is the use of MRF models, because they allow to model in a natural way image pixels contextual information. In particular, we consider a GMRF model, generalised for a complex valued N-dimensional random vector, $\boldsymbol{X}=\boldsymbol{X}_{\text{Re}}+j\boldsymbol{X}_{\text{Im}}$, where $\boldsymbol{X}_{\text{Re}}=[\boldsymbol{X}_1^{\text{Re}}\ \boldsymbol{X}_2^{\text{Re}}\ \ldots\ \boldsymbol{X}_N^{\text{Re}}]^{\text{T}}$ and $\boldsymbol{X}_{\text{Im}}=[\boldsymbol{X}_1^{\text{Im}}\ \boldsymbol{X}_2^{\text{Im}}\ \ldots\ \boldsymbol{X}_N^{\text{Im}}]^{\text{T}}$ are real valued vectors. If we assume real and imaginary parts to be statistically independent, and adopt two hyperparameters, β_R and β_I, for real and imaginary part of the image, respectively, the a-priori joint pdf is:

$$\begin{aligned} f_{\boldsymbol{X}}(\boldsymbol{x};\beta_R,\beta_I) &= f_{\boldsymbol{X}_{\text{Re}}}(\boldsymbol{x}_{\text{Re}};\beta_R)f_{\boldsymbol{X}_{\text{Im}}}(\boldsymbol{x}_{\text{Im}};\beta_I)= \\ &=\frac{1}{Z_R}\exp\{-U(\boldsymbol{x}_{\text{Re}},\beta_R)\}\frac{1}{Z_I}\exp\{-U(\boldsymbol{x}_{\text{Im}},\beta_I)\}= \\ &=\frac{1}{Z_R}\exp\left\{-\frac{1}{2\beta_R^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i}\left(x_i^{\text{Re}}-x_j^{\text{Re}}\right)^2\right\}\times \\ &\times\frac{1}{Z_I}\exp\left\{-\frac{1}{2\beta_I^2}\sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i}\left(x_i^{\text{Im}}-x_j^{\text{Im}}\right)^2\right\}, \end{aligned} \tag{7}$$

where Z_R and Z_I are the partition functions [7], $\mathcal{N}=\{\mathcal{N}_i, i=1,..,N\}$ is a neighbourhood system of second order, (consisting of the nearest 8 pixels of pixel i), and N is the number of pixels. Substituting (7) in (6), we get:

$$\hat{\boldsymbol{x}}_{MAP}=\arg\min_{\boldsymbol{x}\in\Omega}\left[\frac{\|\mathbf{A}(\boldsymbol{x})-\boldsymbol{y}\|^2}{2\sigma_n^2}+U(\boldsymbol{x}_{\text{Re}},\beta_R)+U(\boldsymbol{x}_{\text{Im}},\beta_I)\right] \tag{8}$$

Since $\|\mathbf{A}(\boldsymbol{x})\text{-}\boldsymbol{y}\|^2$ is a non-convex functional, it is attractive to use a convex a priori model (GMRF), so that the risk of being trapped in local minima is reduced. In fact, the GMRF a priori term is a sum, over all the pairs of neighbouring pixels, of quadratic potentials, and so it can help against false solutions. The minimization in (8) has been performed using a deterministic and effective algorithm, the coniugate-gradient method. This method suffers of the presence of local minima in the functional (8); however, the use of a quadratic regularization term, and the estimation of the hyperparameters via a robust statistical algorithm, as explained in the next section, make the overall procedure very robust respect to this aspect, as confirmed by the results presented in section 3.

2.2 Regularization Parameters Estimation Procedure

Hyperparameters estimation can be classified in estimation from *complete data* and estimation from *incomplete data* [6]. Estimation from complete data consists of computing hyperparameter value by direct observation of image $\boldsymbol{x}$, while estimation from incomplete data occurs when we try to achieve it starting from the data $\boldsymbol{y}$ (actually observed data), linked to the image $\boldsymbol{x}$ through a non linear mapping. Of course, the only available is estimation from incomplete data.

In order to implement estimation scheme (8), we have to estimate the value for hyperparameters β_R and β_I from available data set. For this aim, a very popular strategy, used for ML estimation from incomplete data, is the Expectation-Maximisation (EM) algorithm [6]. For the case of a unique hyperparameter, it solves the ML estimation problem repeating two steps, an expectation and a maximisation step, until convergence:

STEP E (expectation):

$$Q\left(\beta \mid \beta^{(k)}\right) = E\left[\ln f_X(X;\beta) \mid Y = y, \beta^{(k)}\right] ;$$

STEP M (maximisation):

$$\beta^{(k+1)} = \arg\max_{\beta} Q\left(\beta \mid \beta^{(k)}\right).$$

The convergence is guaranteed at least to a local optimum. In our case, the estimation problem is separable respect to the two quantities to be estimated, and the EM solution, at the $(k+1)$-*th* iteration, read [6]:

$$\begin{aligned} \left(\beta_R^{(k+1)}\right)^2 &= \frac{E\left[U(\boldsymbol{X}_{\mathrm{Re}},\beta_R) \mid \boldsymbol{Y} = \boldsymbol{y}, \beta_R^{(k)}\right]}{N} \\ \left(\beta_I^{(k+1)}\right)^2 &= \frac{E\left[U(\boldsymbol{X}_{\mathrm{Im}},\beta_I) \mid \boldsymbol{Y} = \boldsymbol{y}, \beta_I^{(k)}\right]}{N} \end{aligned} \tag{9}$$

The analytical evaluation of the expectations in the last is not possible, because of non-linearity, so we have to approximate them by sampling the a posteriori distribution, which is a MRF distribution [6], too. We use the Metropolis algorithm [7] for this generation, which is very easy to implement also in presence of non-linearity, but it is very time consuming. However, we use a fast implementation of the algorithm, based on similar considerations to those presented in [1]. We argue that this statistical algorithm is robust respect to non-linearity, in the sense that it generally produces "good" samples, and usually is able to get out from an eventual local-minima of the distribution, producing a hyperparameter estimation which is not dependent from local-minima solutions. In the next section, we substantiate such considerations via experimental data inversions; exhaustive analysis of the method via simulated data has been already presented in literature.

3. EXPERIMENTAL ANALYSIS

We investigate the performance on two experimental data sets, usually known as the "Marseille" data, provided from the CCRM laboratories in Marseille, France [3]. The measurement set-up configuration is suitable for our method, because a 2-D measurement system with vertical cylinder target is considered. The data relate to dielectric targets composed of one or two filled dielectric cylinders with radius a = 15mm. These cylinders are placed about 30mm from the azimuthal positioner axis. The estimation of the relative permittivity of the objects, made by experimenters [3] via other traditional measurement methods, has resulted in a value of $\varepsilon_r = 3 \pm 0.3$ (i.e., $\chi = 2 \pm 0.3$).

3.1 Case 1: Single Dielectric Object

The first case considered is related to a single dielectric cylinder. We consider a subset of the available data, the ones related to experiments conducted using incident fields in the range 4-8 GHz. We consider a domain under investigation 18×18 cm^2 wide, larger than the minimum region required to enclose the scatterer. In all the reconstructions, we used a given minimal spatial resolution, forced from the sampling rule given in Ref. [1], since we have noticed than a further increase of the resolution does not improve the reconstruction quality.

After the incomplete data estimation of the two hyperparameter's value (one for the real part and one for the imaginary part of the complex image), such estimated values have been used in the MAP procedures to reconstruct the unknown contrast profile [8].

Data at 4 GHz have been inverted using the background (absence of object) as starting point; instead, data at 5-8 GHz have been inverted exploiting a "frequency hopping" technique [9] where the initial point at a given frequency was the reconstruction at the immediately lower frequency.

Fig.1 represents the MAP reconstructions obtained using data at various frequencies, both for real and imaginary part of the image, on the left and to right, respectively.

The result of the reconstruction searching 42×42 unknowns using the incident field at 4 GHz is shown in Fig.1(a). As it is shown, the estimated contrast maximum value ($\chi \approx 2.35$) is slightly higher than the expected value. This notwithstanding, it is possible to recover the location and the size of the object. The result of this reconstruction is used as initialisation point for the following reconstruction at 5 GHz, shown in Fig.1(b) searching 42×42 unknowns for the contrast function. In this case, the estimated contrast maximum value is $\chi \approx 2.5$. The main differences between the two reconstructions are the diameter of the estimated real part of the target, which is smaller for the 5 GHz case, and a less oscillating imaginary part (for the 5 GHz case).

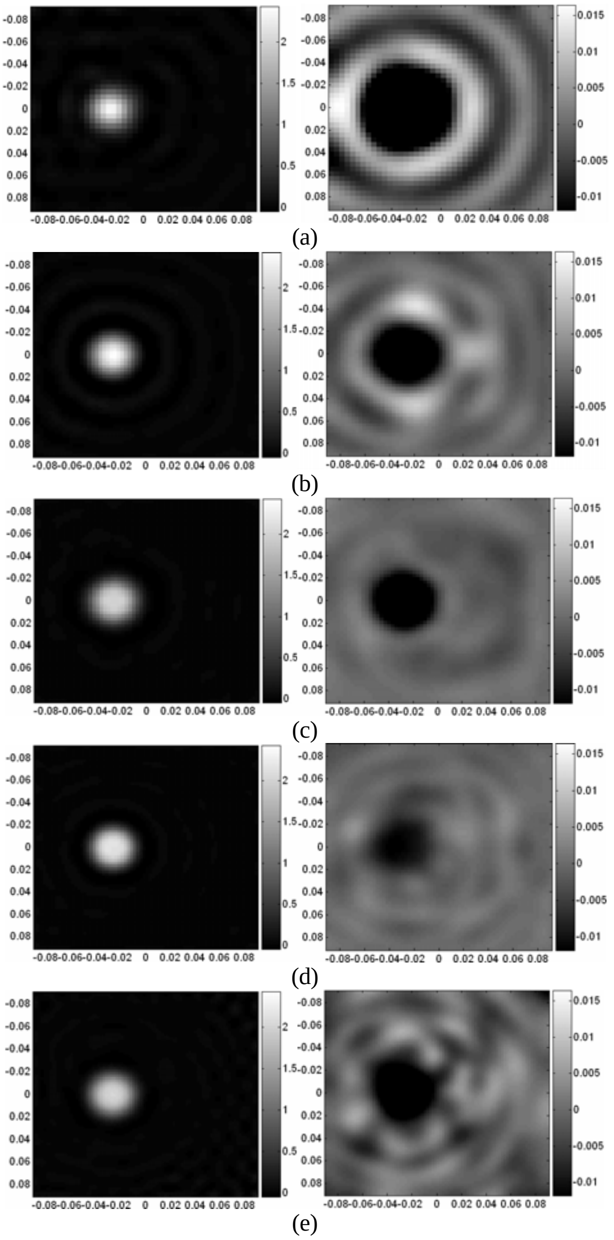

Figure 1 - For all image, left: real part, right: imaginary part. (a) Reconstruction obtained at 4 GHz; (b) reconstruction obtained at 5 GHz, using as starting point the reconstruction at 4 GHz; (c) reconstruction obtained at 6 GHz, using as starting point the reconstruction at 5 GHz; (d) reconstruction obtained at 7 GHz, using as starting point the reconstruction at 6 GHz; (e) reconstruction obtained at 8 GHz, using as starting point the reconstruction at 7 GHz.

A significant improvement can be observed at 6 GHz, searching for 52×52 unknowns (see Fig.1(c)). The reconstruction is clearly more accurate than the ones obtained at 4 and 5 GHz.

Proceeding by hopping [9] from 7 GHz to 8 GHz, and increasing the number of unknowns from 62×62 to 72×72, we refine upon the quality of the reconstructed permittivity profile. As final result, the maximum value of the estimated contrast is $\chi \approx 2$.

As a final comparison, the same image cut of the real part of the reconstructions are presented in Fig.2; this shows the improvements that are obtained by using an increasing working frequency.

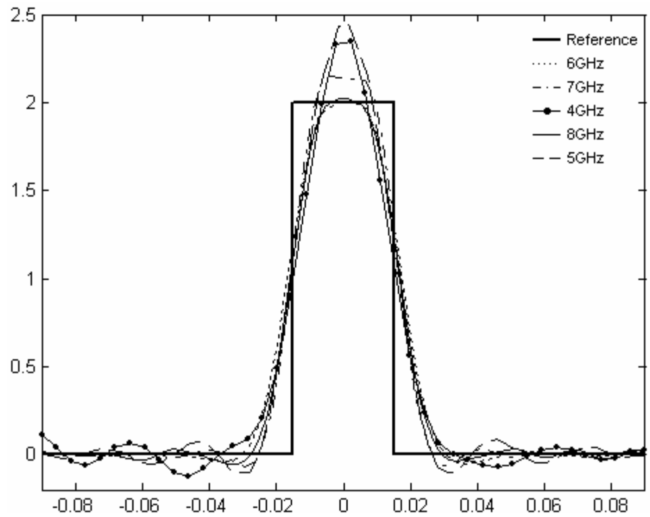

Figure 2 - Cuts of nominal and reconstructed profiles of Fig.1.

3.2 Case 2: Double Dielectric Object

We consider as second case the double dielectric cylinder, with data collected in the range 4-8 GHz. As in section 3.1, we choose a region under test 18×18 cm^2 wide and we use the background as starting point for the 4 GHz case, and the same frequency hopping technique at 5-8 GHz described in the previous sub-section.

Similarly to the previous reconstruction experiment, we first perform estimation from incomplete data of the two hyperparameter's value. After, we perform the MAP reconstructions obtained at the various frequencies (Fig. 3), both for real part and for imaginary part of the image, on the left and the right, respectively.

It is interesting to observe in Fig.3, the effectiveness of the procedure to separating the two objects, also using the lowest frequency. The quality of the reconstructions is clearly improved as the working frequency increases. As a final comparison, the same image cut of the real part of the reconstructions are presented in Fig. 4.

4. CONCLUSIONS AND FUTURE WORK

Statistical based regularization represents a valid approach to circumvent the problems of classic regularization techniques, especially regarding the ability in avoiding a manual and supervised setting of the regularization parameters. The presented algorithm is tested on a well-known experimental data set, providing good quality solutions, avoiding local-minima, and remaining totally unsupervised. The reconstructed shapes of the objects and their permittivity values are an effective representation of the real ones, especially if compared with the results of other approaches [3].

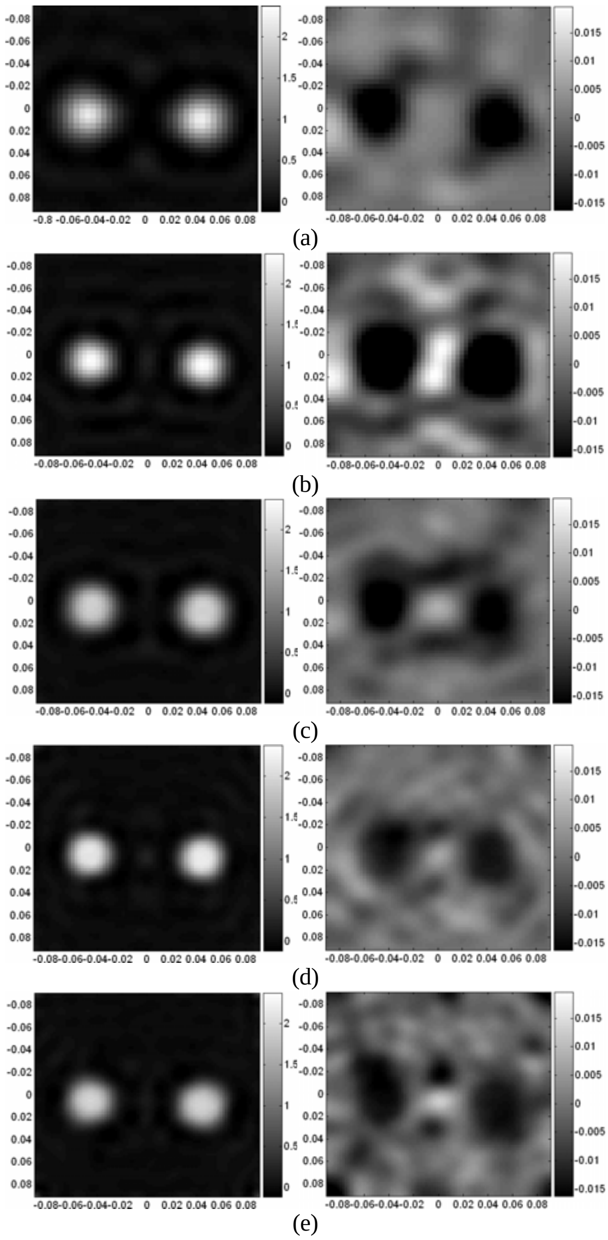

Figure 3 - For all image, left: real part, right: imaginary part. (a) Reconstruction obtained at 4 GHz; (b) reconstruction obtained at 5 GHz, using as starting point the reconstruction at 4 GHz; (c) reconstruction obtained at 6 GHz, using as starting point the reconstruction at 5 GHz; (d) reconstruction obtained at 7 GHz, using as starting point the reconstruction at 6 GHz; (e) reconstruction obtained at 8 GHz, using as starting point the reconstruction at 7 GHz.

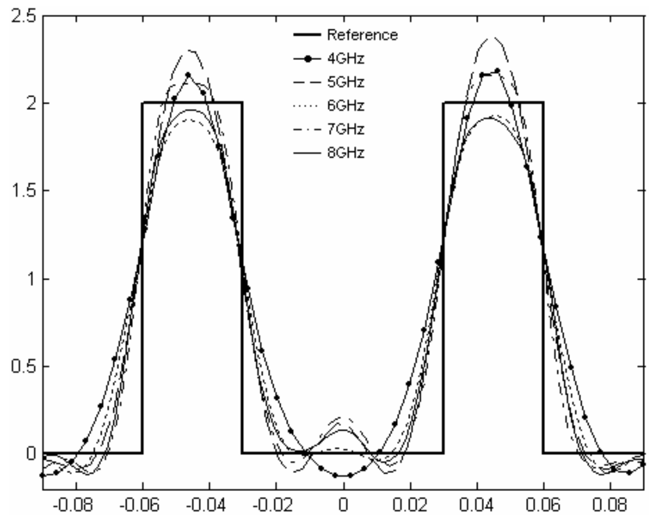

Figure 4 - Cuts of nominal and reconstructed profiles of Fig.2.

These facts confirm what already deduced via simulation and theoretic analysis in previous papers about the approach [4,5,8,9]. Further works are concerned with the experimentation with other available real data sets.

REFERENCES

[1] L. Garnero, A. Franchois, J. P. Hugonin, Ch. Pichot, N. Joachimowicz, "Microwave Imaging – Complex Permittivity Reconstruction by Simulated Annealing", *IEEE Trans. Micr. Th. Tech.*, vol. 39, pp. 1801-1807, 1991.

[2] G. Demoment, "Image Reconstruction and Restoration. Overview of Common Estimation Structures and Problems", *IEEE Trans. Sign. Proc.*, vol. 12, pp. 2024-2036, 1989.

[3] "Special Section: Testing Inversion Algorithms against Experimental data", K. Belkebir, M. Saillard Eds., *Inverse Problems*, vol. 17, pp. 1565-1571, 2001.

[4] T. Isernia, V. Pascazio, R. Pierri, "A Nonlinear Estimation Method in Tomographic Imaging", *IEEE Trans. Geosci. Rem. Sensing*, vol. 35, pp. 910-923, 1997.

[5] V. Pascazio, G. Ferraiuolo, "Statistical Regularization in Linearized Microwave Imaging through MRF-Based MAP Estimation: Hyperparameter Estimation and Image Computation", *IEEE Trans. Image Proc.*, vol. 12, pp. 572-582, 2003.

[6] S. S. Saquib, C. A. Bouman, K. Sauer, "ML parameter estimation for Markov random fields with applications to Bayesian tomography," *IEEE Trans. Image Processing*, vol. 7, pp. 1029–1044, July 1998.

[7] S. Z. Li, *Markov Random Field Modeling In Image Analysis*, Springer, 2001.

[8] G. Ferraiuolo, V. Pascazio, G. Schirinzi, "Maximum a Posteriori Estimation of Height Profiles in InSAR Imaging", *IEEE Geosci. Rem. Sensing Letters*, vol. 1, pp. 66-70, 2004.

[9] O. Bucci, L. Crocco, T. Isernia, V. Pascazio, "Inverse Scattering Problems with Multi-Frequency Data: Reconstruction Capabilities and Solution Strategies", *IEEE Trans. Geosci. Rem. Sensing*., vol. 38, pp. 1749-1756, 2000.