Numerical solution of a nonlinear least squares problem in digital breast tomosynthesis

G. Landi, E. Loli Piccolomini, J. G. Nagy

1Department of Mathematics, University of Bologna
2Department of Mathematics and Computer Science, Emory University
E-mail: germana.landi@unibo.it

Abstract. In digital tomosynthesis imaging, multiple projections of an object are obtained along a small range of different incident angles in order to reconstruct a pseudo-3D representation (i.e., a set of 2D slices) of the object. In this paper we describe some mathematical models for polyenergetic digital breast tomosynthesis image reconstruction that explicitly take into account various materials composing the object and the polyenergetic nature of the x-ray beam. A polyenergetic model helps to reduce beam hardening artifacts, but the disadvantage is that it requires solving a large-scale nonlinear ill-posed inverse problem. We formulate the image reconstruction process (i.e., the method to solve the ill-posed inverse problem) in a nonlinear least squares framework, and use a Levenberg-Marquardt scheme to solve it. Some implementation details are discussed, and numerical experiments are provided to illustrate the performance of the methods.

1. Introduction
Digital tomosynthesis is an emerging technique in digital radiography with the aim of observing a 3D object with low dose radiation [1, 2]. It acquires 2D image projections of a 3D object for varying incident angles. From the set of the 2D acquired images, the reconstruction algorithms should be able to reconstruct any number of slices of the 3D object. In order to make the exam less invasive, the angular range is limited, which requires special consideration in the reconstruction algorithms. For this reason, iterative algorithms are preferable to standard Filtered Backprojection.

In this paper, we consider a reconstruction model taking into account the polyenergetic nature of the x-ray beam and the different materials composing the object. In [3] the authors present a breast tomosynthesis polyenergetic model but they consider only the percentage of one material, while, in [4], the polyenergetic model is extended to the multi-material case by considering the linear attenuation coefficients of each material. The results reported in [4] show that the multimaterial polyenergetic model produces images of improved quality with respect to those obtained by the classical FBP method, reducing noise and beam hardening artifacts. In the following, we present a multi-material polyenergetic model based on the mass attenuation coefficients of the materials composing the object and we reconstruct pseudo-3D images for the mass coefficients of each material. The linear attenuation coefficient of a material represents its attenuation property but it linearly depends on the density (for example, the vapor water and the liquid water have different linear attenuation coefficient values), while the mass attenuation coefficient of a material represents its attenuation property per unit mass.
For this reason, the mass attenuation coefficients have great interest in clinical radiography. In particular, in breast tomosynthesis, the object is composed of air, glandular tissue, adipose tissue and microcalcifications.

The noise on the projection data is a mixture of Poisson noise, coming from x-ray scattering, and Gaussian noise introduced in the digital registration. Since it is reported in [5] that the quality of the reconstructed image is not severely influenced by the noise model, in this paper we suppose to have only Gaussian noise.

The resulting nonlinear least squares problem is large size and ill-posed. Hence, regularization is necessary to compute a good solution. We apply the Levenberg-Marquardt (LM) method that, in our experiments, has shown good performance in terms of reconstructed image quality and computational time. The numerical results presented here for a simulated breast tomosynthesis imaging problem show that the approach is promising and it should be further investigated with more numerical tests on real data with different numerical methods.

In sections 2 and 3 we describe the considered multi-material polyenergetic model and the LM method, respectively. In section 4 we provide some numerical results to illustrate the performance of the numerical methods, and in 5, we present some conclusions.

2. The reconstruction model
Traditionally, the reconstruction models assume, for simplicity, that the source ray is monochromatic, even if the x-ray photons are emitted at different energies in a defined range. Ignoring the energy dependence in the model produces the so called “beam hardening” artifacts in the reconstructed image. In [3] the authors proposed a model accounting for a polyenergetic x-ray. If we consider a 3D object discretized in \(N_v\) voxels, a system detector with \(N_p\) pixels moving through \(N_{\theta}\) angles, the projection in the pixel \(i\) for the incident angle \(\theta\) is expressed as:

\[
b_i^\theta = \sum_{e=1}^{N_e} s_e \exp \left( - \sum_{j=1}^{N_v} a_{i,j}^\theta \mu_{j,e} \right) + \eta_i^\theta, \quad i = 1, \ldots, N_p, \quad \theta = 1, \ldots, N_{\theta},
\]

where \(s_e\) is the term accounting for the energy fluency, i.e. the product of x-ray energy with the number of incident photons at that energy with \(N_e\) discrete energy levels, \(a_{i,j}^\theta\) is the length of the incident ray passing through the voxel \(j\) and contributing to pixel \(i\), \(\mu_{j,e}\) is the linear attenuation coefficient in voxel \(j\) at the energy \(e\) and \(\eta_i^\theta\) represents measurement errors.

In [4] the model (1) was modified by supposing the object constituted of \(N_m\) different materials, each with a specific linear attenuation coefficient \(c_{m,e}\), so that the the linear attenuation coefficients \(\mu_{j,e}\) can be approximated as:

\[
\mu_{j,e} = \sum_{m=1}^{N_m} w_{j,m} c_{m,e} \quad \text{with} \quad \sum_{m=1}^{N_m} w_{j,m} = 1.
\]

The unknowns \(w_{j,m}\) are the weights of the linear attenuation coefficients \(c_{m,e}\) of each material and they depend on the percentage of the \(m\)th material in voxel \(j\) at the x-ray energy \(e\). Hence equation (1) can be written in terms of the \(w_{j,m}\) as:

\[
b_i^\theta = \sum_{e=1}^{N_e} s_e \exp \left( - \sum_{j=1}^{N_v} a_{i,j}^\theta \sum_{m=1}^{N_m} w_{j,m} c_{m,e} \right) + \eta_i^\theta, \quad i = 1, \ldots, N_p, \quad \theta = 1, \ldots, N_{\theta}.
\]

In the following, we consider the mass attenuation coefficients \(\delta_{j,e}\) for the voxel \(j\) at the x-ray energy \(e\) that are related to the linear attenuation coefficients \(\mu_{j,e}\) by:

\[
\mu_{j,e} = \rho_j \delta_{j,e}
\]
where \( \rho_j \) is the density of the composite materials in voxel \( j \). Similarly to equation (2), we can write \( \delta_{j,e} \) as a linear combination of the mass attenuation coefficients \( d_{m,e} \) of the \( m \)-th material at the x-ray energy \( e \) through some weights \( \tilde{w}_{j,m} \):

\[
\delta_{j,e} = \sum_{m=1}^{N_m} \tilde{w}_{j,m} d_{m,e} \quad \text{with} \quad \sum_{m=1}^{N_m} \tilde{w}_{j,m} = 1.
\] (5)

Moreover, the densities \( \rho_j \) can be expressed as:

\[
\rho_j = \frac{1}{\sum_{m=1}^{N_m} \tilde{w}_{j,m} r_m}
\] (6)

where \( r_m \) is the density of the \( m \)-th material. By substituting (4), (5) and (6) in (1) we obtain the expression of the projection in pixel \( i \) at the angle \( \theta \) as a function of the mass attenuation coefficients:

\[
b_i^\theta = \sum_{e=1}^{N_e} s_e \exp \left( -\sum_{j=1}^{N_v} a_{i,j}^\theta \sum_{m=1}^{N_m} \tilde{w}_{j,m} d_{m,e} \frac{1}{\sum_{m=1}^{N_m} \tilde{w}_{j,m} r_m} \right) + \eta_i^\theta, \quad i = 1, \ldots N_p, \quad \theta = 1, \ldots N_\theta.
\] (7)

If we introduce the \( N_v \times N_m \) matrix \( \tilde{W} \) with entries \( \tilde{w}_{j,m} \), the \( N_p N_\theta \times N_v \) matrix \( A \) with entries \( a_{i,j}^\theta \), the vectors \( \mathbf{b} \) and \( \mathbf{\eta} \) of length \( N_p N_\theta \) obtained by storing the projections \( b_i^\theta \) and the noise elements \( \eta_i^\theta \), respectively, and the vector \( s \) of length \( N_e \) containing the values \( s_e \) then, we can write equation (7) in matrix-vector form as:

\[
\bar{r}(\tilde{W}) = \mathbf{0}
\] (8)

where \( \bar{r} \) is a function from \( \mathbb{R}^{N_v N_m} \) to \( \mathbb{R}^{N_p N_\theta} \) defined as

\[
\bar{r}(\tilde{W}) := \mathbf{b} - \left( \exp \left( -A \left( \frac{\tilde{W} D}{\bar{W} R} \right) \right) \mathbf{s} + \mathbf{\eta} \right), \quad \mathbf{D}_{m,e} := d_{m,e}, \quad \mathbf{R}_{m,e} := \frac{1}{r_m}
\] (9)

where \( \mathbf{D} \) and \( \mathbf{R} \) have size \( N_m \times N_e \). The exp(\cdot) function and the division operation are intended to be performed element-wise, while the arguments of the exponential function is computed with matrix-matrix and matrix-vector operations. In practice, the x-ray energies can be estimated by using the well-known spectra models [6], the matrix \( \mathbf{D} \) and \( \mathbf{R} \) can be obtained by taking x-ray transmission measurements of objects that have known dimension, density and material composition, while the noise mean and variance can be estimated through a calibration process. The unknown weights \( \tilde{w}_{j,m} \) of the mass attenuation coefficients can be obtained by solving the inverse problem (8).

3. Numerical solution of the inverse problem

For the numerical solution of the inverse problem (8), we employ the widely used least squares approach requiring to solve the constrained problem:

\[
\min_{\tilde{W}} \frac{1}{2} \| \bar{r}(\tilde{W}) \|^2
\]

s.t. \( \tilde{W} 1_{N_m} = 1_{N_e} \)
where \( \mathbf{1}_N \) is the vector of length \( N \) whose elements are all 1. This isn’t the only possible approach for solving the inverse problem (8). For example, a Bayesian approach could be considered to model the uncertainty in the estimated matrices \( \mathbf{D} \) and \( \mathbf{R} \) or to strengthen the Poissonian component of the noise.

Upon variable substitution for \( \tilde{w}_{j,1} = 1 - \sum_{m=2}^{N_m} \tilde{w}_{j,m} \), problem (10) becomes the unconstrained nonlinear least squares problem

\[
\min_{\mathbf{W}} \frac{1}{2} \| \mathbf{r}(\mathbf{W}) \|^2
\]

where, using Matlab-like notation, \( \mathbf{W} = \mathbf{W}(:, 2 : N_m) \) and

\[
\mathbf{r}(\mathbf{W}) := \mathbf{b} - \left( \exp \left( -\mathbf{A} \left( \mathbf{W}^{-1}(\cdot, m), \begin{bmatrix} \mathbf{D} \\ \mathbf{R} \end{bmatrix} \right) \right) s + \eta \right).
\]

Usually, problem (11) is ill-posed and the Jacobian matrix \( \mathbf{J} \) of the residual vector \( \mathbf{r}(\mathbf{W}) \) is severely ill-conditioned. In order to account for the ill-conditioning of \( \mathbf{J} \), for the numerical solution of (11), we consider a globally convergent modified LM method [7] defined by the general iteration:

\[
\mathbf{W}_{k+1} = \mathbf{W}_k + \alpha_k \mathbf{p}_k
\]

where \( \mathbf{p}_k \) satisfies

\[
(\mathbf{J}^T_k \mathbf{J}_k + \lambda_k \mathbf{I}) \mathbf{p}_k = -\mathbf{J}_k^T \mathbf{r}_k
\]

and the step-length \( \alpha_k \) is determined by an Armijo-like procedure with initial unit step-length [7]. The scalar \( \lambda_k \) is chosen according to the following criterion:

\[
\lambda_k = \max(c_1, \min(c_2, \| \mathbf{J}_k^T \mathbf{r}_k \|)), \quad \lambda_0 = c_2
\]

with \( 0 < c_1 < \lambda_k < c_2 \). This criterion prevents severe ill-conditioning in the matrix \( \mathbf{J}^T_k \mathbf{J}_k + \lambda_k \mathbf{I} \) and accelerates convergence when the residuals are small. Observe that the proposed criterion follows the original Levenberg-Marquardt strategy of selecting \( \lambda_k \) directly instead of using a trust-region approach to manipulate the choice of \( \lambda_k \). Global convergence is ensured by the line-search procedure.

4. Numerical results

In this section, we present some preliminary results obtained for a simulated breast imaging problem. The experiments were performed using Matlab 7.14 (R2012a) on a Sun Fire V40z server consisting of four 2.4GHz AMD Opteron 850 processors with 16GB RAM. In our test problem, we considered one simulated 3D phantom object of size \( N_v = 64 \times 64 \times 7 \), made of four ellipses consisting of a tissue mixture with varying percentages of glandular and adipose tissue and with homogeneous background made of a mixture of 50% adipose and 50% glandular tissue. The central slice of the exact 3D weights image of the mass attenuation coefficients of the glandular material is shown in figure 1. Observe that the weights of the mass attenuation coefficients of the adipose material can be obtained from those of the glandular material by using the equation \( \sum_{m=1}^{N_m} \tilde{w}_{j,m} = 1 \). The matrix \( \mathbf{A} \) was obtained by considering \( N_\theta = 15 \) equispaced projection angles from \(-17^\circ\) to \(17^\circ\) and, for each projection angle, the ray trace matrix was computed using a cone beam model with Siddon’s algorithm for ray tracing [8]. Finally, \( N_e = 37 \) different levels of energy were considered from 10 keV to 28 keV, in 0.5 keV steps. Gaussian noise was added to the 15 projection images to simulate experimental errors.

We define the noise level, \( nl \), as the ratio between \( \| \eta \|_2 \) and \( \| \mathbf{b} - \eta \|_2 \): the noise levels \( 10^{-3} \) and \( 10^{-4} \) were used in our experiments. We note that lower is the noise level, lower is the noise on the image; sometimes the Signal to Noise ratio (SNR) is used to measure the noise intensity that,
conversely, is higher when the noise is greater. The nonlinear least squares problem (11) was solved by the LM method (13)-(14) where central finite differences were used to approximate the Jacobian matrix $J_k$ and the values of the parameters $c_1$ and $c_2$ were respectively fixed at $10^{-2}$ and $10^3$ for $nl = 10^{-3}$, and $10^{-2}$ and 1 for $nl = 10^{-4}$. The linear system (14) was solved by using the Cholesky decomposition. The LM method exhibits a semi-convergent behavior [9] and the iteration plays the role of the regularizing parameter. Since, in this work, we are mainly interested in analyzing the capability of the reconstruction approach (11) to compute the weights of the mass attenuation coefficients, in figure 1, we present the central slice of the restored 3D weights images of the glandular mass attenuation coefficients for the glandular material with the lowest relative error. Figure 2 depicts the horizontal line slice through the center of the weights images of figure 1. For $nl = 10^{-3}$, the optimal error value 0.0408 was obtained at the 9th iteration; for $nl = 10^{-4}$, the optimal error value 0.0139 was achieved at the 9th iteration. Figure 3 shows the restored weights images in the YZ plane and figure 4 depicts the corresponding line slice through their center. Finally, figure 5 presents the iteration history of the relative error of the unknown weights $W_k$ (left) and the objective function (right) for $nl = 10^{-3}$. The plot on the left illustrates the semi-convergence property of the LM method, while the plot on the right shows that the objective function becomes nearly flat in the neighborhood of the iteration where the lowest relative error value is obtained. This behavior has been observed in all the performed numerical tests.

Figure 1: Central slice of the restored 3D weights images of the mass attenuation coefficients for the glandular material. From left to right: exact slice, reconstruction slice for $nl = 10^{-3}$ and for $nl = 10^{-4}$.

Figure 2: Horizontal profile through the center of the central slice of the 3D weights images of the mass attenuation coefficients for the glandular material for $nl = 10^{-3}$ (left) and $nl = 10^{-4}$ (right).

5. Conclusions
In this work we have proposed an approach to digital breast tomosynthesis where information about both the different materials composing the object and the different energy levels of the
Figure 3: Central slice of the glandular weights images in the YZ plane. From left to right: exact slice, reconstruction slice for $nl = 10^{-3}$ and $nl = 10^{-4}$.

Figure 4: Profile through the center of the central slice of the 3D weights images of the mass attenuation coefficients for the glandular material in the YZ plane for $nl = 10^{-3}$ (left) and $nl = 10^{-4}$ (right).

Figure 5: Objective function (left) and relative error (right) histories for $nl = 10^{-3}$.

x-ray beam is incorporated in the reconstruction model. In particular, we have examined the inverse problem of determining the mass attenuation coefficients of the various materials at the different x-ray energy levels and we have proposed a Levenberg-Marquardt-type method for its solution. Regularization has been enforced by early termination of the iteration. Initial results of simulated test problems show the potential of the proposed approach, since the method well separate the two considered materials and it seems robust with respect to the noise, although there is still research to be done regarding the regularized inversion method. Moreover, we think that the first step should be to develop a parallel implementation of the proposed approach in order to be able to handle large-scale clinical data with thousands of unknowns.

References
[1] Grant D 1972 IEEE Trans. Biom. Eng. 19 20–28
[2] Gobbins J and Godfrey D 2003 Phys. Med. Biol. 48 R65–R106
[3] Chung J, Nagy J and Sechopoulos I 2010 SIAM J. Imag. Sci. 3 133
[4] Bustamante V, Nagy J, Feng S and Sechopoulos I 2013 SIAM J. Scient. Comput. 35 S192–S208
[5] Lu Y, Chan H, Wei J, Goodsitt M, Carson P, Hadjiiski L, Schmitz A, Eberhard J and Claus B 2006 Med. Phys. 33 3781–3795
[6] Cranley K, Gilmore B, Fogarty G and Despords L 1997 Catalogue of diagnostic x-ray data and other data Tech. Rep. 78. Institute of Physics and Engineering in Medicine (IPEM) York
[7] Nocedal J and Wright S J 2006 Numerical Optimization (New York: Springer)
[8] Siddon R L 1985 Med. Phys. 12 252–255
[9] Kaltenbacher B, Neubauer A and Scherzer O, 2008 Iterative Regularization Methods for Nonlinear Ill-Posed Problems (Walter de Gruyter, Berlin)