Research Article

Journal of Algorithms & Computational Technology

Journal of Algorithms & Computational Technology
0(0) 1–13

DOI: 10.1177/1748301816668022
act.sagepub.com

SAGE

A fast Total Variation-based iterative algorithm for digital breast tomosynthesis image reconstruction

E Loli Piccolomini[1] and E Morotti[2]

Abstract

In this work, we propose a fast iterative algorithm for the reconstruction of digital breast tomosynthesis images. The algorithm solves a regularization problem, expressed as the minimization of the sum of a least-squares term and a weighted smoothed version of the Total Variation regularization function. We use a Fixed Point method for the solution of the minimization problem, requiring the solution of a linear system at each iteration, whose coefficient matrix is a positive definite approximation of the Hessian of the objective function. We propose an efficient implementation of the algorithm, where the linear system is solved by a truncated Conjugate Gradient method. We compare the Fixed Point implementation with a fast first order method such as the Scaled Gradient Projection method, that does not require any linear system solution. Numerical experiments on a breast phantom widely used in tomographic simulations show that both the methods recover microcalcifications very fast while the Fixed Point is more efficient in detecting masses, when more time is available for the algorithm execution.

Keywords

Digital breast tomosynthesis imaging, Fixed Point method, iterative regularization algorithm, Total Variation regularization, tomographic images reconstruction

Date received: 30 October 2015; accepted: 15 March 2016

Introduction

Digital breast tomosynthesis (DBT) is a 3D emerging technique for the diagnosis of breast tumors that has some advantages over the traditional 2D mammography suffering from the fact that the lesions can be hidden by overlaying tissues in the plane representation of a 3D object. In DBT, the breast volume is reconstructed in a stack of 2D slices and the structure is resolved in space, reducing the impact of the overlapping tissues on the tumor and making easier the tumor detection by the radiologist.

The volume image is reconstructed from 2D cone beam projections data acquired at a limited number of views over a limited angular range, in order to reduce the X-ray dose across the body. Due to the data incompleteness, DBT image reconstruction is challenging. One pass algorithms, such as Filtered Back Projection (FBP), traditionally used in complete data tomography, introduce artifacts and noise and lose the information about tissue density when used in limited angle tomography.[1–3] Iterative statistical reconstruction algorithms, that maximize the similarity between the computed and measured projections at each iteration and enable the introduction of priors, have some advantages over FBP in the noise reduction and in the identification of the object borders. Investigation of iterative algorithms derived from complete data tomography and applied to DBT, such as Maximum Likelihood or Algebraic Iterative Algorithms (ART), can be found in Wu et al.[4] and Zhang et al.[5] Mathematically, they can be formulated as minimization problems of the form:

$$\min_f \mathcal{J}(f, M, g) \tag{1}$$

[1]Department of Mathematics, University of Bologna, Bologna, Italy
[2]Department of Mathematics, University of Padova, Padua, Italy

Corresponding author:
E Loli Piccolomini, Department of Mathematics, University of Bologna, Bologna, Italy.
Email: elena.loli@unibo.it

where f is the image to be reconstructed, M is the projection operator and g is the data vector. In this paper, we consider the function $\mathcal{J}(f, M, g)$ as the least squares function:

$$\mathcal{J}(f, M, g) = \|Mf - g\|_2^2 \tag{2}$$

As evidenced in Lu et al.,[2] tomosynthesis reconstruction is an illposed problem both for the ill posedness of the continuous projection operator[6] and for the data incompleteness, that causes infinite solutions to the underdetermined least squares problem (2).[2] For this reason, the idea of using a regularization operator has been introduced in Lu et al.[2] and Sidky et al.,[7] with the double purpose of loosening up the consistency with the data and of selecting an image with the prescribed regularity among the infinite possible solutions. The regularization function makes assumption on the reconstructed image and forces the choice of one of the infinite possible solutions.

Recently, the compressed sensing theory[8] has been used in Computed Tomography. If the image is supposed sparse in some domain, the minimization of the 1-norm in that domain guarantees the sparsity of the solution. For breast images, where the interest is to identify microcalcifications and/or masses, the image gradient is supposed to be sparse and the Total Variation (TV) regularization function has been successfully employed.[1,7,9–12]

In this case, the minimization problem can have a constrained formulation:

$$\min_f TV(f) \quad s.t. \quad \|Mf - g\|_2^2 \leq \sigma \tag{3}$$

or an equivalently unconstrained formulation:

$$\min_f \|Mf - g\|_2^2 + \lambda TV(f). \tag{4}$$

Some iterative algorithms for the solution of problem (3) or (4) applied to the reconstruction of 3D tomographic images have been investigated in literature.[2,9,10,13–17] All the cited algorithms are first order algorithms, since they use only the gradient information of the objective function. We refer in particular to Jensen et al.[15] for a comparative analysis of first order methods for the minimization of a least-squares TV function applied to 3D tomography.

Aim and contribution

In this paper, we apply to DBT image reconstruction a Newton-like method, the Lagged Diffusivity Fixed Point (FP) algorithm by Vogel,[18] that uses second order information of the TV function. The FP algorithm is efficiently used in other imaging applications, such as deblurring or denoising, and in this paper, we show that it performs very well even in DBT imaging reconstruction. On the basis of the results obtained in Jensen et al.,[15] where first order methods are applied to 3D tomography, we compare the proposed algorithm with the Scaled Gradient Projection (SGP) algorithm accelerated by Barzilai–Borwein rules, as the representative of efficient first order methods for this particular application.

The main contribution of the paper is the efficient implementation of a method with second order information for the reconstruction of DBT images, i.e. a very large size problem with unstructured matrix. The method has better performance when compared with the methods traditionally used for DBT images reconstruction both in the first iterations and at convergence, so that in the final images not only the microcalcifications but also the masses can be better distinguished. Hence, it could be potentially used in clinical applications.

The paper is organized as follows. In the section ''Mathematical model of digital tomosynthesis'' we present the mathematical model of DBT image formation and the numerical optimization problem for the DBT image reconstruction; in the section ''Algorithms for DBT image reconstruction'' we describe, with implementation details for this particular application, the proposed FP method and the SGP method considered for comparison; in the section ''Numerical experiments'' the numerical results on a digital breast phantom are reported and lastly in the section ''Conclusions'' we make some final observations.

Mathematical model of digital tomosynthesis

Digital tomosynthesis

Digital tomosynthesis is a cone-beam tomographic technique where the projections are obtained along a smaller range of incident angles.[19] Usually, 10–25 projections along a range of up to 40–45° are obtained and from these projections a pseudo-3D representation of the object is reconstructed, with lower resolution in the z-direction perpendicular to the detector plane (see Figure 1 for a schematic representation of the tomosynthesis system).[20] There are different geometries of motion of the X-ray tube, but the reconstruction algorithms can be easily adapted to each one without significant changes.[5] In our case, the source rotates in the yz plane with x-coordinate equal to zero. The detector is positioned onto the xy plane and the breast is compressed over it.

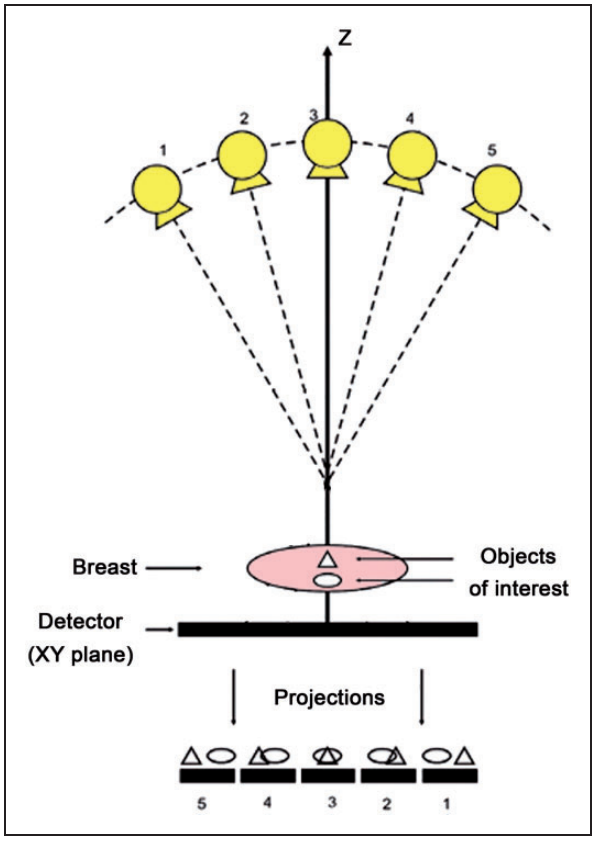

Figure 1. Tomosynthesis system in *yz* plane.

The mathematical model of image formation

If we suppose a monochromatic X-ray source, the continuous mathematical model of the image formation process is described by Beer's law that relates the measured values $\bar{g}_i^{(\theta)}$ at each pixel i of the detector, for a fixed angle θ, with the attenuation coefficient μ for each position $\mathbf{w}$ of the object:

$$\bar{g}_i^{(\theta)} = \exp\left(-\int_{L_\theta} \mu(\mathbf{w})\mathrm{d}l\right) + \bar{\eta}_i^{(\theta)}, \quad i = 1 \dots N_p, \quad \theta = 1, \dots N_\theta \tag{5}$$

where:

- N_p is the number of pixels in the detector (N_p is of the order of millions in the real systems);[20]
- N_θ is the number of angles (common values are $10 \leq N_\theta \leq 25$ in the real systems);[20]
- L_θ is the line followed by the X-ray beam through the object;
- $\mu(\mathbf{w})$ is the linear attenuation coefficient, at the position $\mathbf{w} = (x, y, z)$, depending on the material in the object and characterizing the structures inside the object. We remind that denser materials, such as calcifications, have a greater attenuation coefficient;
- $\bar{\eta}_i^{(\theta)}$ is the noise measured at the detector (pixel i, angle θ) and it includes scattering and electronic noise.

The discretization of equation (5) is:

$$\bar{g}_i^{(\theta)} = \exp\left(-\sum_{l=1}^{N_v} m_{il}^{(\theta)} \mu_l\right) + \bar{\eta}_i^{(\theta)} \tag{6}$$

where:

- N_v is the number of voxels (few billions in the real systems) in the discretized 3D object;
- $m_{il}^{(\theta)}$ is the element of a matrix $M^{(\theta)}$ obtained with the ray-tracing technique proposed by Siddon.[21]

If we take the negative logarithm of (6) and we reorder all the resulting projections and noise elements in vectors g and η of length $N_p \cdot N_\theta$, we obtain the matrix equation:

$$g = Mf + \eta \tag{7}$$

where the matrix $M = m_{il}^{(\theta)}$ is sparse with size $(N_p \cdot N_\theta) \cdot N_v$, and f is the vector with elements $\mu_l, l = 1, \dots N_v$.

Numerical model for tomosynthesis image reconstruction

DBT images are the grayscale representation of the attenuation coefficients of each voxel of the breast. Hence, aim of the reconstruction algorithms is to compute the values of the vector f, given the projection data g.

From the previous model (6), it is clear that the problem of reconstructing f can be reduced to the solution of the linear system (7). In the real cases $N_\theta \cdot N_p < N_v$, hence the system is underdetermined and it has infinite possible solutions. Moreover, since the system comes from the discretization of the ill-posed integral equation (5),[22] the noise is amplified in the computed solution. For this reasons, a prior information is used embedded in a regularization term, in order to choose a solution with the characteristics of the true object and to prevent the noise amplification.

Therefore, the reconstructed image f can be computed as the solution of a minimization problem of the form:

$$\min_f J(f) = \frac{1}{2}\|Mf - g\|_2^2 + \lambda R(f) \tag{8}$$

where $R(f)$ is a suitable regularization function and $\lambda > 0$ is the regularization parameter. In tomographic reconstructions a widely used regularization function is

the TV function[18,23] defined as:

$$TV(x) = \int_{\Omega} |\nabla x| \mathrm{d}x \tag{9}$$

where Ω is the image domain and $|\cdot|$ is the $L2$ norm.

Since the TV is not differentiable, usually it is substituted by a smooth differentiable function. For example, in Jensen et al.[15] the Huber function is used instead of the TV. In deblurring and denoising applications (see for example[18,24,25]) a small parameter β is often added and a smoothed version of the TV function is obtained as:

$$TV_{\beta}(f) = \int_{\Omega} |\nabla f + \beta| \mathrm{d}f \tag{10}$$

with β a small positive value. In this paper, we use this last strategy and $TV_{\beta}(f)$ is considered in place of $TV(f)$ in the following.

Hence, the reconstructed image is the solution of the minimization problem:

$$\min_{f} \mathcal{J}(f) = \frac{1}{2}\|Mf - g\|_2^2 + \lambda TV_{\beta}(f) \tag{11}$$

Since the reconstructed image f should have nonnegative values, the constraint $f \geq 0$ can be added to the problem (11).

Algorithms for DBT image reconstruction

In this section, we briefly revise the algorithms that we compare for the solution of problem (11): the SGP and the FP methods. The SGP method finds a nonnegative solution of the minimization problem (11). It is a gradient-like method[26] accelerated by using a scaling matrix and the Barzilai–Borwein rules for the choice of the steplength. This algorithm has been recently proposed in imaging applications with very good results in terms of efficiency and precision.[27,28] The FP method is a Newton-like method for the solution of the minimization problem (11),[26] whose descent step is computed by solving a linear system. The coefficient matrix is an approximation of the Hessian of the objective function $\mathcal{J}(f)$. The FP algorithm has been proposed by Vogel[18] and Vogel and Oman[25] for the solution of image denoising and image deblurring TV regularized problems due to its fast convergence. It has never been used in tomographic applications, probably because the solution of a linear system for each FP iteration is considered too expensive in terms of time and mainly of memory requirements for a very large size application.

In this paper, we present an efficient implementation of the FP method tailored for the tomographic problem; at the same time the FP method maintains its desirable convergence properties.

Scaled Gradient Projection method

The SGP method is a gradient-like method solving the constrained minimization problem:

$$\begin{aligned} \min \quad & \mathcal{J}(f) = \frac{1}{2}\|Mf - g\|_2^2 + \lambda TV_{\beta}(f) \\ \text{s.t.} \quad & f \geq 0 \end{aligned} \tag{12}$$

exploiting scaling strategies and step-length selection rules to improve the convergence speed (see Bonettini et al.[27] and references therein). In each iteration, the scaling matrix is a diagonal matrix $D_k = \mathrm{diag}\,(d_1^{(k)}, d_2^{(k)}, \ldots, d_{N_v}^{(k)})$, whose entries are computed effortless as $d_i^{(k)} = \min\{L, \max\{\frac{1}{L}, (f_k)_i\}\}\ \forall i = 1, \ldots, N_v$, and L is an appropriate threshold. In addition, starting from the well-known Barzilai–Borwein original ideas, in Bonettini et al.[27] the authors propose an efficient strategy for the step-length updating rule: given a suitable positive range $[\alpha_{\min}, \alpha_{\max}]$ in input, we define α_k choosing alternatively a value between

$$\alpha_k^{(1)} = \frac{s_{k-1}^t D_k^{-1} D_k^{-1} s_{k-1}}{s_{k-1}^t D_k^{-1} z_{k-1}} \quad \text{and} \quad \alpha_k^{(2)} = \frac{s_{k-1}^t D_k z_{k-1}}{z_{k-1}^t D_k D_k z_{k-1}} \tag{13}$$

where $s_{k-1} = f_k - f_{k-1}$ and $z_{k-1} = \nabla J\mathcal{J}(f_k) - \nabla \mathcal{J}J(f_{k-1})$.

Avoiding to introduce significant computational costs, simple projections on $\Omega := \{f : f \geq 0\}$ are added to the basic model (11).

A trace of the algorithm is reported in Table 1.

Implementation notes. The step 5 of the SGP algorithm computes the most appropriate step length μ_k for the descent passage, with a backtracking loop. This is the most expensive point of the algorithm, because it requires one function evaluation in each inner execution plus one matrix-vector product.

For the evaluation of $\mathcal{J}(f)$, the $TV_{\beta}(f)$ is computed by discretizing (10) as:

$$\sum_{i=1}^{N_v^x} \sum_{j=1}^{N_v^y} \sum_{l=1}^{N_v^z} \sqrt{(\nabla_x F_{ijl})^2 + (\nabla_y F_{ijl})^2 + (\nabla_z F_{ijl})^2 + \beta^2} \tag{14}$$

where N_v^x, N_v^y, N_v^z are the number of voxels along the x, y, z axis respectively and F_{ijl} is the three dimensional matrix of size $N_v^x \times N_v^y \times N_v^z$ obtained by reshaping the vector f of length N_v. The computation of $\nabla_x F_{ijl}, \nabla_y F_{ijl}$ and $\nabla_z F_{ijl}$ is made by forward differences.

Moreover, only one function assessment and one gradient evaluation are involved for each main iteration k.

Table 1. The SGP algorithm. For more details, see [27].

SGP algorithm to solve: $J(f) = \frac{1}{2}\|Mf - g\|_2^2 + \lambda TV_\beta(f)$, *s.t.* $f \in \Omega$
choose initial f_0 set parameters $\gamma \in (0,1), 0 < \alpha_{\min} < \alpha_{\max}$, $K \in \mathbb{N}$, $\beta_{bt} > 0$ set k = 0 while (convergence) 1. choose $\alpha_k \in [\alpha_{\min}, \alpha_{\max}]$ and the scaling matrix D_k with the BB rules (13) 2. compute the projection $h_k = P_{\Omega D_k^{-1}}(f_k - \alpha_k D_k \nabla J(f_k))$ 3. if $h_k = f_k$ then stop 4. compute the descent direction $d_k = h_k - f_k$ 5. backtracking loop to compute the steplength μ_k: set $\mu_k = 1, f_{\max} = max_{0 \le j \le \min(kK-1)} J(f_{k-j})$ while $(J(f_k + \mu_k d_k) > f_{\max} + \beta_{bt}\lambda_k \nabla J(f_k)^t d_k)$ $\mu_k = \gamma \mu_k$ end 6. $f_{k+1} = f_k + \mu_k d_k$ 7. $k = k + 1$ end

SGP: Scaled Gradient Projection.

We use a convergence criterium based on the difference between two successive iterates, since in imaging applications it is useless to go on with the iterations if the image does not change any more:

$$\|f_k - f_{k-1}\|_2 < tol_{SGP}$$

Lagged Diffusivity FP method

The FP method is a Newton-like method for the solution of the minimization problem (11), with an FP method that makes use of approximated second order information.

To solve the convex optimization problem (11), the first order condition

$$\nabla \mathcal{J}(f) = M^t(Mf - g) + \lambda \nabla TV_\beta(f) = 0 \qquad (15)$$

must be satisfied and the proposed FP method, given a starting guess f_0, computes the sequence

$$f_{k+1} = (M^t M + \lambda L(f_k))^{-1} M^t g, \quad \forall k = 0, 1, \ldots.$$

that leads to the following updating rule:

$$f_{k+1} = f_k - (M^t M + \lambda L(f_k))^{-1} \nabla J(f_k) \qquad (16)$$

In this notation, $L(f_k)$ is the discretization matrix of the diffusion operator $L(f)$ acting on f as $L(f)f = \nabla TV_\beta(f)$ as described in Vogel.[18]

Table 2. The FP algorithm.

FP algorithm to solve: $J(f) = \frac{1}{2}\|Mf - g\|_2^2 + \lambda TV_\beta(f)$
choose initial f_0 k = 0 while (convergence) 1. compute the gradient $g_k = \nabla J(f_k)$ 2. compute the approximated hessian $H_k = M^t M + \lambda L(f_k)$ 3. solve the linear system $H_k s_k = -g_k$ with CG method 4. backtracking loop to compute the steplength α_k 5. $f_{k+1} = f_k + \alpha_k * s_k$ 6. k = k + 1 end

FP: Fixed Point.

Following the notations used in Vogel,[18] $L(f)$ can be written as:

$$L(f) = D_x^t diag(\psi'(f)) D_x + D_y^t diag(\psi'(f)) D_y + D_z^t diag(\psi'(f)) D_z \qquad (17)$$

where D_x is the matrix of the backward differences for the first order operator in the x direction (the same is for D_y and D_z in the y and z direction, respectively), and

$$\psi(f) = 2\sqrt{f^2 + \beta^2}$$

A scheme of the FP algorithm for the solution of the problem (11) is shown in Table 2.

Implementation notes. In step 1, $\nabla J(f_k)$ is computed as:

$$\nabla J(f_k) = M^t(Mf_k - g) + \lambda L(f_k) f_k$$

In step 2, the matrix H_k is never stored since the matrix L is made of only five diagonal vectors (of maximum length N_v) and M_k^{tMf} is computed by two matrix-vector products.

Step 3 is the most computationally expensive part, because a large size linear system is to be solved with the CG algorithm. We require the CG not to converge exactly, in order to get more regularized final solutions. For this reason, the CG algorithm is stopped when the residual 2-norm is smaller than a tolerance of 10^{-3} or a maximum number of iterations is reached.

Finally, step 4 performs a backtracking procedure, as in the SGP algorithm, to compute the step-length α_k.

The FP stopping criterion is set as:

$$\|f_k - f_{k-1}\|_2 < tol_{FP}$$

Numerical experiments

The experiments of this section aim to show the efficiency and effectiveness of the FP method, comparing it

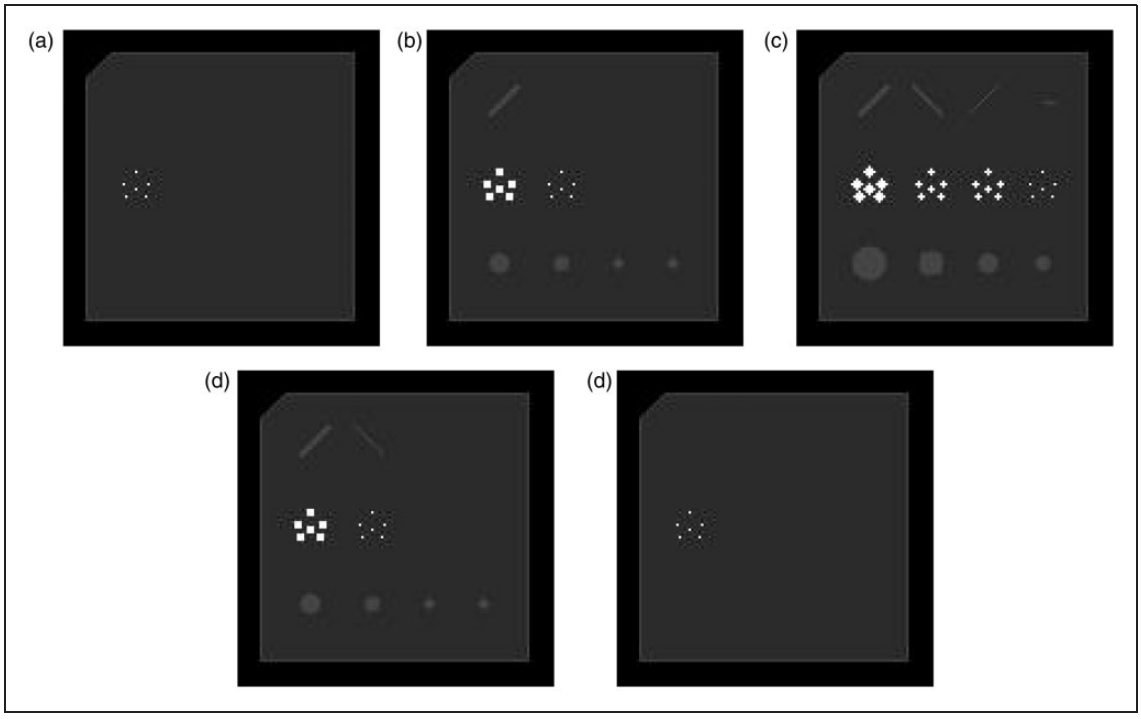

Figure 2. Central layers of the digital mammographic phantom CIRS mod. 015, sliced in 15 digital planes with resolution of 128×128 pixels. (a–e): Layers (6–10).

with the SGP method. We analyze the results obtained by the two methods after few iterations of the algorithms, i.e. what can be obtained during clinical trials, and at convergence, when a posteriori executions can be performed offline. We think that both the simulations are interesting in practical use.

All the methods are implemented in Matlab R2015a and the experiments are performed on a computer equipped with two processors Intel Pentium G2120 3.1 GHz, with 3.6 GB RAM, using Linux operating system.

Table 3. Attenuation coefficients of different materials, related to a 20 keV scan.

Object	Attenuation coefficient
Adipose tissue	0.1703
Skin	0.24
Fibers and masses	0.27
Microcalcifications	3

Test problem

The tests are made on a digital version of a real accreditation phantom, called "CIRS mod. 015."[29] In our tests, it is discretized in $128 \times 128 \times 15$ voxels with a resolution of $1 \times 1 \times 3\,\text{mm}^3$. These dimensions are proportional to a real case.

Figure 2 shows the central slices of this phantom: inside a voxel-thin boundary of simulated skin, objects like fibers, microcalcification and masses are neatly put in a uniform background of adipose tissue. Outside the skin-made boundary, air is simulated with null attenuation coefficient and it is part of the volume we want to reconstruct together with the phantom. This phantom is used to check the competency of mammographic systems because the included objects are very important for the early detection of a breast cancer. In more detail, they are of different dimensions and thickness, in order to analyze and compare the 3D graphic resolution of different reconstructions, and their attenuation values are performed taking into account a low X-ray scan of 20 keV.

In Table 3, we report the attenuation coefficients used to create our digital phantom.[30] The central layer includes one-voxel thin microcalcifications on the right, in order to analyze the algorithm even on the smallest objects. In Figure 3, the reverse gray-scale representation of Figure 2(c) is reported to better appreciate the low contrast between structures and adipose background.

In our tests, the detector is made of 128×128 pixels and its extension is the same of the volume area on the xy-plane. Moreover the volume leans on the detector, so it captures all the projections of the CIRS even from the most angled views (since it is smaller than the

volume, because of the surrounding air ring). The tomosynthesis acquisition parameters that we use in our simulations are inspired by the real systems produced by IMS (Internazionale Medico Scientifico). We consider 13 angles, uniformly distributed from -17 and $+17^\circ$. The central cone beam starts 64 cm above the detector and the X-ray monochrome source wheels along the x-direction, with a 59 cm ray.

The simulation problem is defined by computing the projections:

$$g = Mf + \eta$$

where η is Gaussian noise of *relative noise level*:

$$rnl = \frac{\|\eta\|}{\|g\|} = 10^{-3}$$

Evaluation criteria. Outputs are evaluated by computing the *relative error*

$$RE = \frac{\|f_{out} - f^*\|}{\|f^*\|}$$

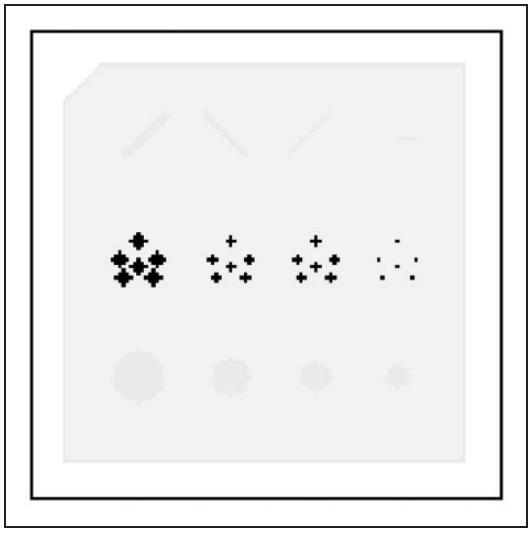

Figure 3. Layer 8, in reverse gray-scale.

where f^* is the true object and f_{out} is the reconstructed image.

Figure 4 shows profiles and objects, on the eighth layer, that we use to analyze all the outputs in more details. Denoising, in fact, is estimated taking the standard deviation value (*StdDev*) on a rectangle in the upper-right side of the central layer, where the actual phantom is uniform (Figure 4(a)): smaller values mean stronger noise reduction. To better highlight the object reconstruction, during the code executions, we analyze some early outputs comparing profiles taken on microcalcifications and on masses along the horizontal lines in Figure 4(b) and (c). In order to analyze the 3D accuracy of the algorithm we focus on the two voxels pointed by arrows in Figure 4(d).

Algorithms parameters. The TV makes use of $\beta = 10^{-6}$ for its differentiability.

The regularization parameter has been heuristically chosen for both algorithms as $\lambda = 0.01$: this choice ensures good results since we carefully evaluated the robustness of both methods.

Both the algorithms always start from an initial null vector $f_0 = 0$ and the convergence tolerances are $tol_{SGP} = tol_{FP} = 10^{-4}$. A further control is added to ensure the convergence of the solution of problem (11) to the minimum relative error: the algorithms may stop if semiconvergence is achieved with respect to the exact solution f^*, that is

$$\frac{\|f_k - f^*\|}{\|f^*\|} > \frac{\|f_{k-1} - f^*\|}{\|f^*\|}$$

where f^* is the exact image and f_k is the kth computed iteration.

In the SGP code, we used $L = 10$ to construct the scaling matrices D_k, $\alpha_{\min} = 10^{-10}, \alpha_{\max} = 10^5$, and $\alpha_0 = 10^{-2}$ to ensure α_k positivity; we set $K = 0$, $\theta = 0.4$, and $\beta_{bt} = 10^{-4}$ for line-search backtracking loop. We heuristically tested different sets of parameters and we reported the best performing one; the values are the same proposed in Bonettini et al.[27]

The maximum number of iterations for the inner CG loop, in the FP algorithm, is 150.

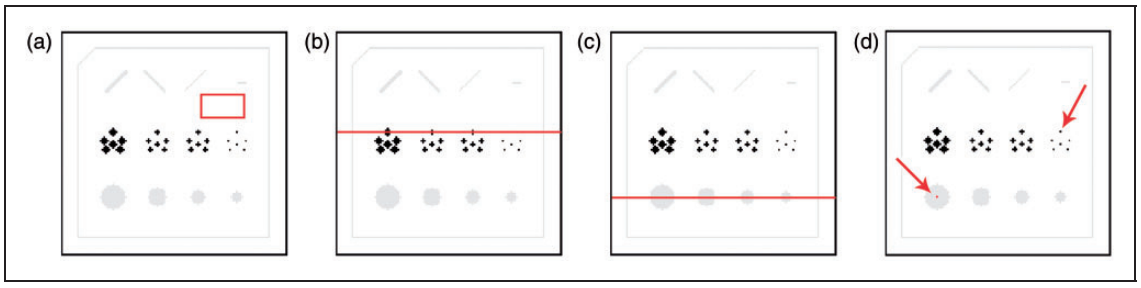

Figure 4. Objects we will focus on to analyze reconstructions deeply.

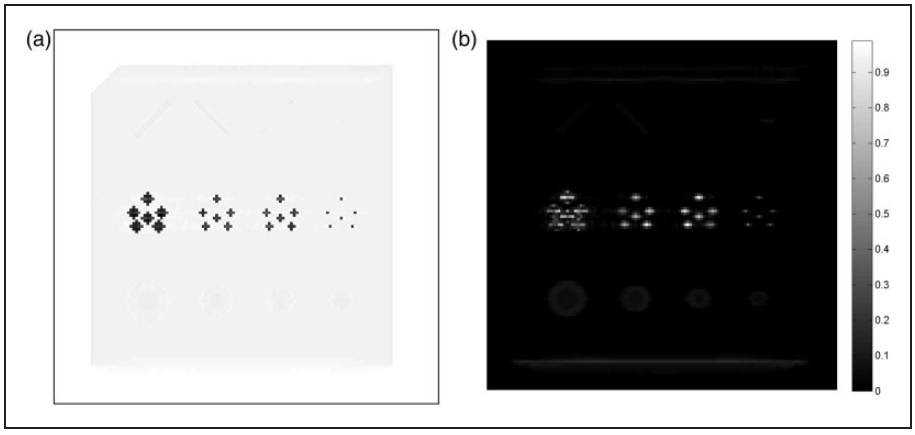

Figure 5. SGP reconstruction in the eighth layer and the related absolute error.
SGP: Scaled Gradient Projection.

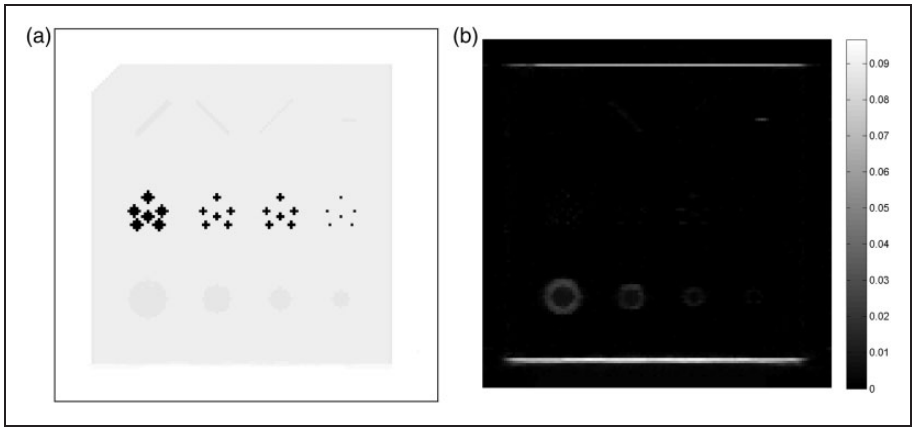

Figure 6. FP reconstruction in the eighth layer and the related absolute error. To notice that the color bar scale is one order smaller than that in Figure 5.
FP: Fixed Point.

Results at the algorithms convergence

Figures 5 and 6 show the central layer (in reverse gray-scale) of SGP and FP reconstructions together with the absolute errors between the original and the reconstructed images. First of all, we can appreciate that all the microcalcifications are detected, even the smallest ones, in both outputs and this is an encouraging remark. Actually speaking, the SGP is faster than the FP method, since it requires about one fifth time to get up to its convergence: more precisely, the SGP convergence takes 105 s while the FP makes use of 505 s to provide its output. On the other hand, the FP output looks much more precise than the SGP output. In fact, focusing on the central layer that has values in [0, 3], the FP pointwise absolute error is always smaller than 10^{-1} and, as we can observe comparing the two color bars in Figures 5 and 6, it is 10 time smaller than the SGP value. In addition, these inaccuracies are mainly located in the CIRS horizontal edges (that are harder to detect precisely, because they are parallel to the X-ray source motion) and around the masses, while only tiny errors are near the microcalcifications. Finally, the FP reconstruction also shows a uniform background, without shadows along the acquisition direction, and this feature is particularly appreciable, because shadows may overlap with other structures and conceal them.

Results after few algorithms iterations

Due to the practical aim, where there is no time to wait for the convergence, it is essential to analyze how the algorithms behave during their execution, with untimely outputs. From now on, we analyze the performance in temporal windows of 5, 20, and 60 s of computations: it is interesting to see how the profiles (related to the lines

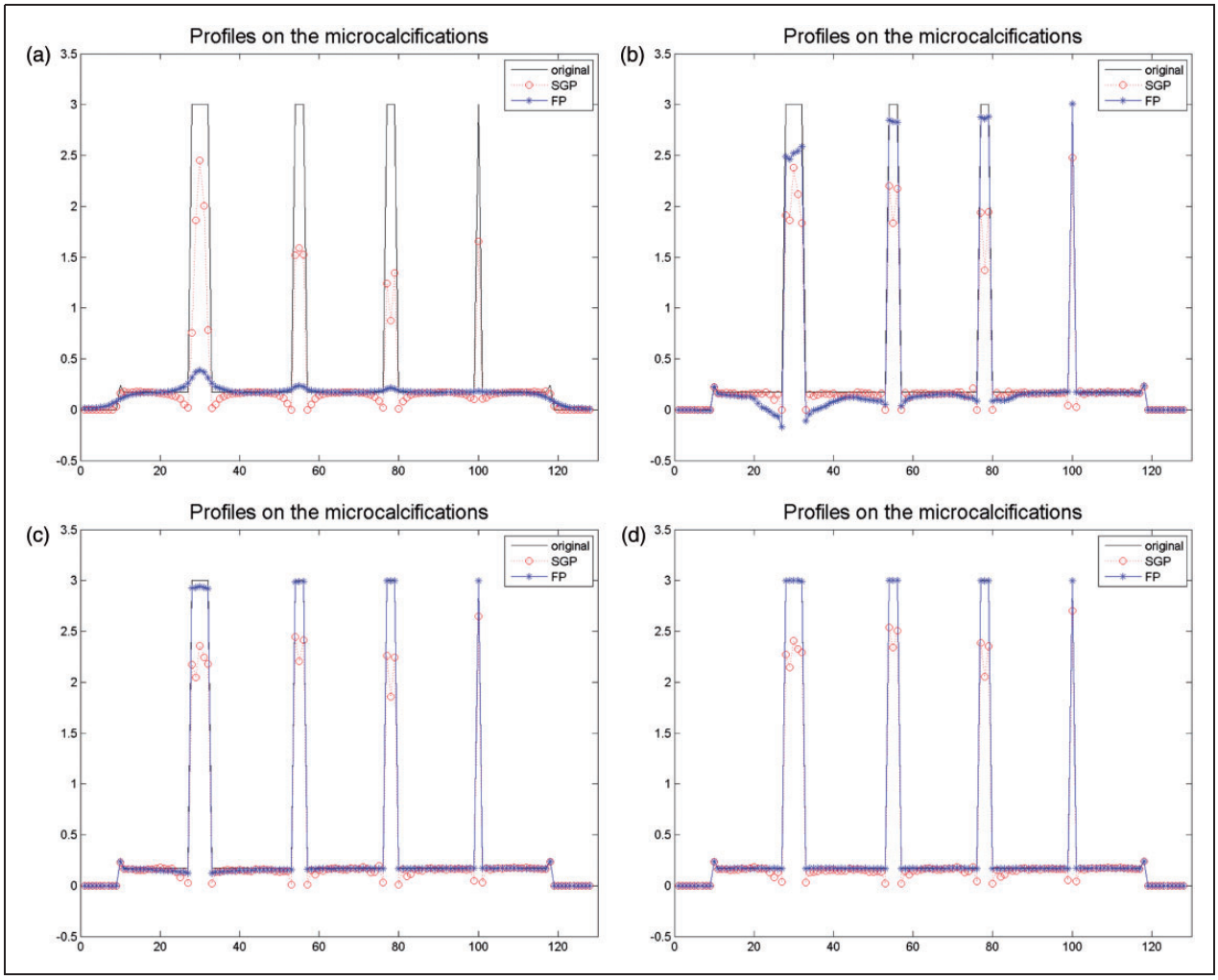

Figure 7. Profiles related to Figure 4(b), taken from some early reconstructions of both methods. In black continuous lines the exact profiles, in red circles the SGP, and in blue asterisks the FP profiles. (a): 5″, (b): 20″, (c): 60″, and (d): Convergence.

plotted in Figure 4(b) and (c)) on the partial reconstructions become closer and closer to their target.

Figures 7 and 8 point out the main differences between the two methods. If we concentrate on the microcalcifications, we can realize how the SGP is really fast in its first computations: it makes the microcalcifications visible after only 5 s, while in that time the FP hardly moves from its initial null guess. On the contrary, in 20 s the FP becomes more effective, since the smallest object detection is very accurate, and later outputs show that the SGP never manage in reaching the tops, even in its convergence. On the other hand, the FP reconstructions are almost perfect.

Focusing on the masses (Figure 8), the gap between the two outputs becomes more and more noticeable since the blue lines representing the FP output approach more and more to their target, as time goes by, while the SGP suffers from inaccuracy when it deals with all these low-attenuation elements.

At last, we cannot appreciate significant improvements between the 60-s and the convergent outputs.

Numerical parameters to compare all the corresponding solutions are reported in Table 4. The rows correspond to the shots at different execution times; 5 s, 20 s, and 60 s. The first column shows the computed relative errors for each reconstruction, taking into account all the layers. It confirms the SGP speed and the FP accuracy: in initial outputs the SGP error is smaller, while for the final reconstructions the FP error is half of the SGP one.

The second column reports the number of performed iterations, together with the number of backtracking executions for the SGP method (between brackets), and the total number of CG inner loops for the FP method (between brackets). As we can see, each FP iteration takes more time than a SGP iteration.

Finally, the third column compares the two algorithms in terms of noise removal through the

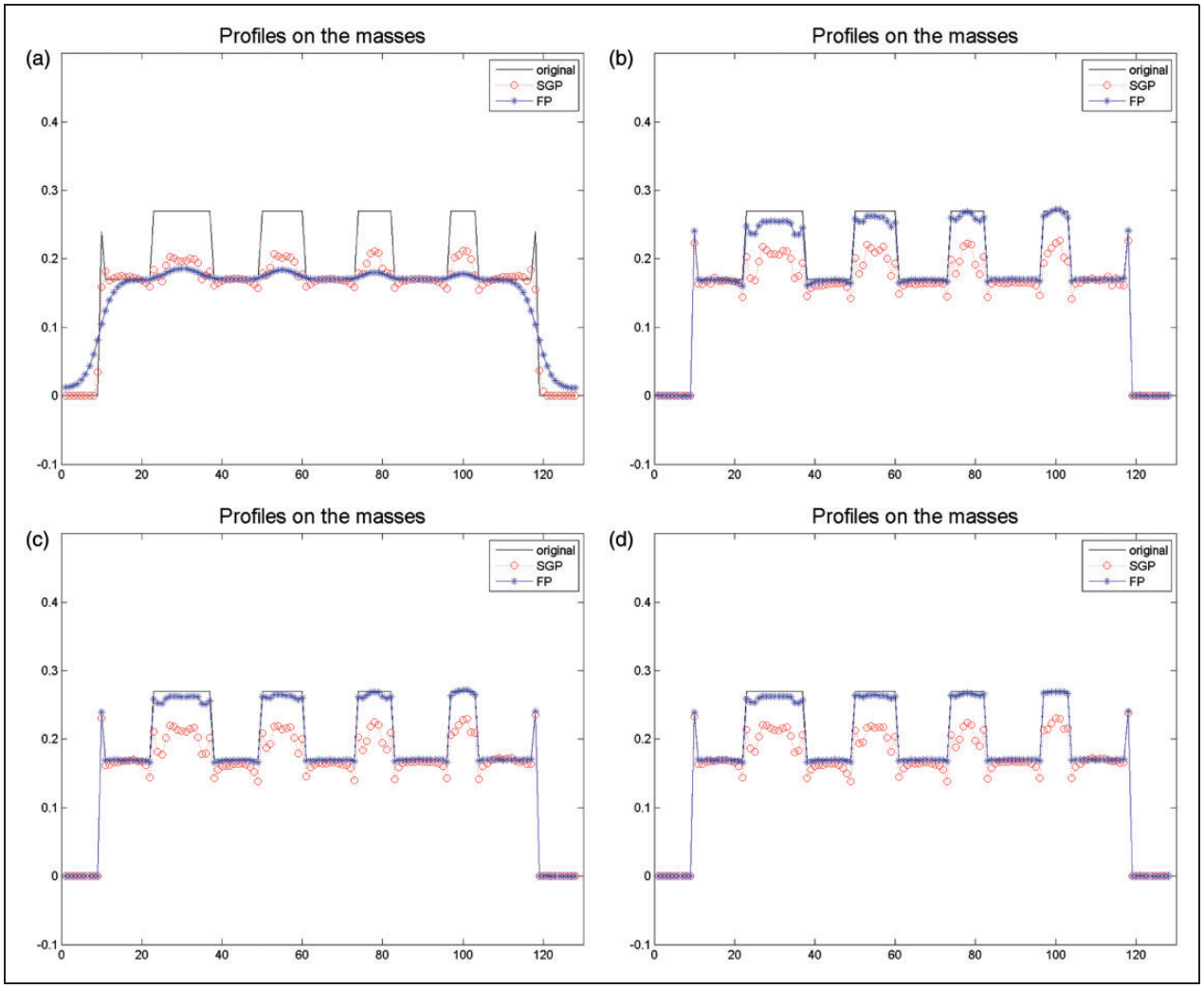

Figure 8. Profiles related to Figure 4(c), taken from some early reconstructions of both methods. In black continuous lines the exact profiles, in red circles the SGP, and in blue asterisks the FP profiles. (a): 5″, (b): 20″, (c): 60″, and (d): Convergence.

Table 4. Results for the FP and SGP methods at different execution times.

	RE		Iterations		Std Dev	
	SGP	FP	SGP	FP	SGP	FP
5 s	0.28135	0.54478	13 (+1)	1 (+85)	1e–3	5.4e–4
20 s	0.17416	0.18484	57 (+15)	4 (+505)	1e–3	3.5e–4
60 s	0.13355	0.10717	174 (+53)	8 (+1105)	8.3e–4	5.6e–4
Convergence	0.12221	0.062893	283 (+84)	62 (+9055)	6.2e–4	3.2e–4

SGP: Scaled Gradient Projection; FP: Fixed Point.

standard deviation value. We remark that the images obtained with the FP method always have a very low value of standard deviation. The values of the means of these test are not reported, since they all approximate the exact value with an error lower than 10^{-3}.

3D accuracy

As we are interested in three-dimensional reconstructions, we now investigate the accuracy of both the methods along the z direction, in the two voxels pointed by arrows in Figure 4(d).

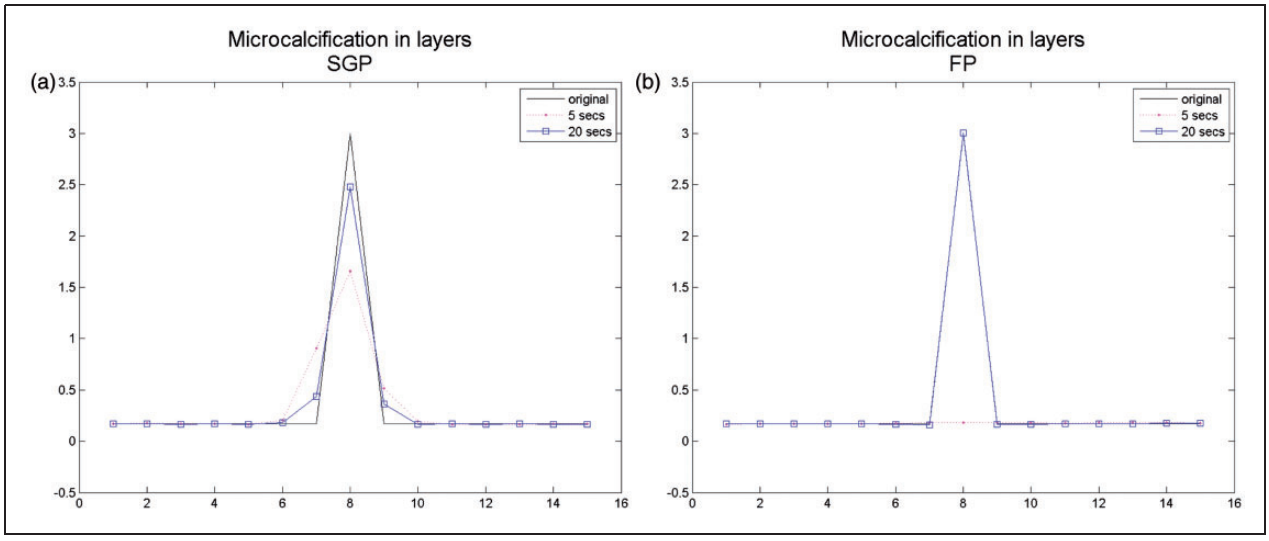

Figure 9. *z*-Profiles related to the one-voxel thin microcalcification: in black continuous lines the exact profiles, in red dots the output after 5 s, and in blue squares output after 20 s. (a): SGP and (b): FP.

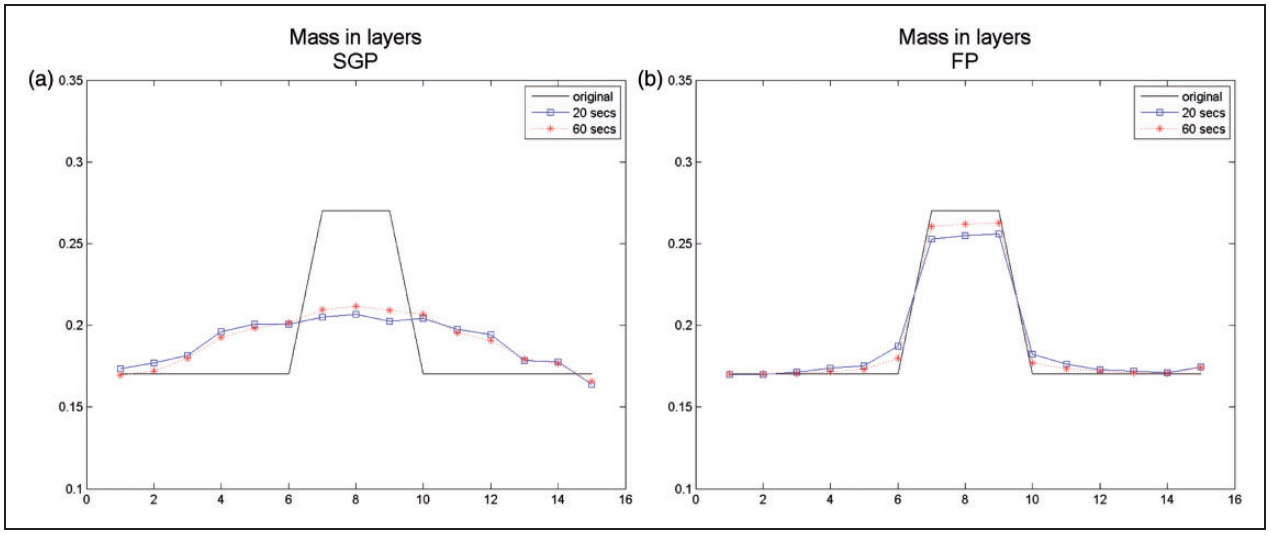

Figure 10. *z*-Profiles related to the central voxel of the biggest mass: in black continuous lines the exact profiles, in blue squares the output after 20 s, and in red asterisks the output after 60 s. (a): SGP and (b): FP.

Looking at the plots in Figure 9 related to a thin microcalcification, we can confirm the excellent capability of both the methods to locate voxel-wise elements of high intensity in their exact positions, even in few iterations. Hence, we plot the values of the reconstructed voxels in the 15 slices along the z axis. This is an important validation for the prefixed model.

On the contrary, the plots in Figure 10 are related to the central voxel of the biggest mass on the left (that is three layers thick) and they reveal the inability of the SGP algorithm to detect lighter and plainer objects; the FP method, on the contrary, needs short time to detect smoother structures. We remark that, for what concerns the masses, in the z direction the two methods produce very different results, since the SGP method detects an object with very low intensity diffused in too many slices, while the FP method accurately finds the intensity and depth of the original structure. These profiles are concerning reconstructions obtained after 20 and 60 s. We notice that if we make many more iterations of the SGP algorithm the relative error does not decrease and the reconstructed image remains the same.

Conclusions

In this paper, we have proposed to use an FP method for the solution of a regularized TV least squares minimization problem for the reconstruction of DBT images. In tomographic applications, only first order methods have been used up to now, because one iteration of a first order method is less expensive than one iteration of a method using approximate second order information. We proposed an efficient implementation of the FP method, suited for DBT applications, that makes it competitive with a fast first order method such as SGP.

Numerical experiments performed on a medium size phantom showed that the FP method is very well performing, both after very few iterations and at convergence. Both the FP and SGP methods detect microcalcifications in a very short time, while in a longer time the FP method recovers different structures, as masses, that cannot be obtained with SGP. For practical purposes, this is very useful when a deeper investigation is needed to make a diagnosis, if the operator has quite a lot of time or a powerful computing configuration available.

Acknowledgements

We thank the I.M.S. Company for the collaboration in the choice of the phantom and in the discussion of the results.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The research is part of the project *Nuovi aspetti della regolarizzazione nell'imaging* funded by the GNCS.

References

1. Kastanis I, Arridge S, et al. 3D digital breast tomosynthesis using Total Variation regularization. *LNCS* 2008; 5116: 621–627.
2. Lu Y, Chan HP, et al. Selective-diffusion regularization for enhancement of microcalcifications in digital breast tomosynthesis reconstruction. *Med Phys* 2010; 37(11): 6003–6014.
3. Jerebko A, Kowarschik M and Mertelmeier T. Regularization parameter selection in maximum a posteriori iterative reconstruction for digital breast tomosynthesis. *LNCS* 2010; 6136: 548–555.
4. Wu T, Moore RH, Rafferty EA, et al. A comparison of reconstruction algorithms for breast tomosynthesis. *Med Phys* 2004; 31(9): 2636–2647.
5. Zhang Y, Chan HP, et al. A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis. *Med Phys* 2006; 33(10): 3781–3795.
6. Kirsch A. *An introduction to the mathematical theory of inverse problems*. New York: Springer, 1996.
7. Sidky EY, Pan X, Reiser IS, et al. Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms. *Med Phys* 2009; 36(11): 4920–4932.
8. Donoho DL. Compressed sensing. *IEEE Trans Inf Theory* 2006; 52: 289–306.
9. Sidky EY, Kao CM and Pan X. Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct. *J X-ray Sci Technol* 2006; 14(2): 119–139.
10. Sidky EY, Jørgensen JH and Pan X. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm. *Phys Med Biol* 2012; 57(10): 3065–3091.
11. Heil U, Frankel S, et al. Total variation regularization in digital breast tomosynthesis: Regularization parameter determination based on small structures segmentation rates. In: *12th fully 3D image reconstruction in radiology and nuclear medicine*, 2013, pp.485–488.
12. Wu J, Mou X and Zhang Y. A fast iterative soft-thresholding algorithm for few-view CT reconstruction. In: *3rd workshop on high performance image reconstruction*, 2011, pp.72–75.
13. Yu H and Wang G. A soft-threshold filtering approach for reconstruction from a limited number of projections. *Phys Med Biol* 2010; 55: 3905–3916.
14. Jørgensen JH, Jensen TL, Hansen PC, et al. Accelerated gradient methods for total-variation-based CT image reconstruction. In: *11th fully 3D image reconstruction in radiology and nuclear medicine*, 2011, pp.435–438.
15. Jensen TL, Jorgensen JH, Hansen PC, et al. Implementation of an optimal first-order method for strongly convex total variation regularization. *BIT Numer Math* 2012; 52: 329–356.
16. Yu H and Wang G. SART-type half-threshold filtering approach for CT reconstruction. *IEEE Access* 2014; 2: 602–612.
17. Ertas M, Yildirim I, Kamasak M, et al. An iterative tomosynthesis reconstruction using total variation combined with non-local means filtering. *BioMed Eng Online* 2014; 13(65): 1–13.
18. Vogel CR. *Computational methods for inverse problems*. Philadelphia, PA, USA: SIAM, 2002.
19. Wu T, et al. Tomographic mammography using a limited number of low-dose cone-beam projection images. *Med Phys* 2003; 30: 365–380.
20. Chung J, Nagy JG and Sechopoulus I. Numerical algorithms for polyenergetic digital breast tomosynthesis reconstruction. *Siam J Imag Sci* 2010; 3(1): 133–152.
21. Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys* 1985; 12(2): 252–255.
22. Hansen PC. *Discrete inverse problems: Insight and algorithms*. SIAM, USA, 2010.
23. Rudin LI, Osher S and Fatemi E. Nonlinear total variation based noise removal algorithms. *Phys D* 1992; 60: 259–268.

24. Zanella R, Boccacci P, Zanni L, et al. Efficient gradient projection methods for edge-preserving removals of Poisson noise. *Inv Probl* 2009; 25(5): 045010.
25. Vogel CR and Oman ME. Iterative methods for total variation denoising. *SIAM J Sci Comput* 1996; 17: 227–238.
26. Nocedal J and Wright SJ. *Numerical optimization*. Springer-Verlag New York, Inc., 1999.
27. Bonettini S, Zanella R and Zanni L. A scaled gradient projection method for constrained image deblurring. *Inv Probl* 2009; 25(1): 015002, 1–23.
28. Benvenuto F, Zanella R, Zanni L, et al. Nonnegative least-squares image deblurring: Improved gradient projection approaches. *Inv Probl* 2010; 26(2): 025004.
29. Computerized imaging reference systems. http://www.cirsinc.com/products/all/47/mammographic-accreditation-phantom/
30. Vedantham S and Karellas A. X-ray phase contrast imaging of the breast: Analysis of tissue simulating materials. *Med Phys* 2013; 40: 041906, 1–8.