3D RECONSTRUCTION OF LOCALIZED OBJECTS FROM RADIOGRAPHS AND BASED ON MULTIRESOLUTION AND SPARSITY

Charles Soussen

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI-CNRS) BP 133, 91403 Orsay Cedex, France

Jérôme Idier

Institut de Recherche en Communications et Cybernétique de Nantes (IRCCyN), BP 92101, 1 rue de la Noë, 44321 Nantes Cedex 3, France

ABSTRACT

We address the reconstruction of a 3D image from a set of incomplete X-ray tomographic data. In the case where the image is composed of one or several objects lying in a *uniform* background, we define a sparse parameterization by considering the active voxels, *i.e.,* the voxels that do not lay inside the background. Estimation of the active voxel densities is performed using the maximum *a posteriori* (MAP) estimator. In order to implement sparse parameter estimation, we design an original multiresolution scheme, which handles coarse to fine resolution images. This scheme affords automatic selection of active voxels at each resolution level, and provides a drastic decrease of the computation time. We finally show the performance of our method on synthetic data.

1. INTRODUCTION

This work deals with the reconstruction of a 3D image from a set of limited X-ray tomographic data. We focus on the case where the image is composed of one or several small size, non uniform objects lying in a uniform background. Without loss of generality, we assign non negative values to the objects and the value 0 to the background. The density function writes:

$$f(x,y,z) \begin{cases} = 0 \text{ if } (x,y,z) \text{ lays inside the background,} \\ > 0 \text{ otherwise.} \end{cases} \quad (1)$$

Such images are often encountered in nondestructive evaluation (NDE) applications, where the objects represent air faults included in a homogeneous metallic material [1, 2].

In the case where the number or angles of projection are limited, classic reconstruction methods aim to discretize the volume $f(x,y,z)$ into a set of cubic voxels $\mathbf{f} = [f_1, \ldots, f_n]^t$ (^{t} denotes vector transposition), and then estimate their values from the X-ray data using the regularization theory, *e.g.,* MAP estimation [1, 3]. Although this approach provides accurate reconstructions, it appears to be consuming, both in memory storage and computation time for high resolution images.

When the image contains a uniform background, we propose to perform an automatic detection of the voxels that lay inside the background, and then only estimate the values of the other voxels. This sparse model is simple, and affords a drastic reduction of the number of parameters, and thus of the reconstruction time. MAP estimation of the voxel densities is performed using iterative local minimization algorithms, for which the final solution and the reconstruction time largely depend on the initial solution. In order to provide an accurate initial solution, and thus to enhance the quality and time of reconstruction, we design a multiresolution scheme, in which the voxel resolution is gradually improved [2, 3]. This scheme includes, at each resolution level r:

1. Detection of the active voxels, from the reconstruction at level $r-1$;
2. Initial estimation of the voxel densities at level r;
3. MAP estimation of the voxel densities at level r.

In the following, we specify the modeling of the data, and the monoresolution reconstruction method (all voxels are active). Then, we introduce sparse parameterization of non negative images, and the corresponding reconstruction method, relying on the multiresolution scheme. We finally evaluate our reconstruction method on a set of limited synthetic data.

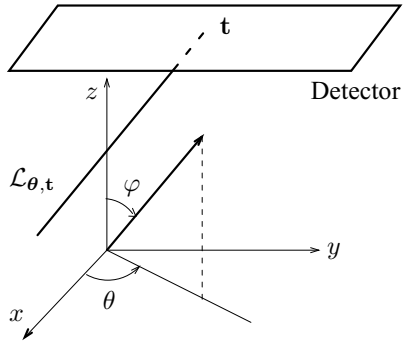

Figure 1: Line projection parameterization: $\boldsymbol{\theta} = (\theta, \varphi) \in [0, 2\pi) \times [0, \pi)$. All detectors $\mathbf{t}$ lay in a single horizontal plane.

2. PROBLEM STATEMENT

2.1. Measurement modeling

The X-ray transform of a 3D image $f(x,y,z)$ is defined by its line integrals:

$$p_{\boldsymbol{\theta}}(\mathbf{t}) = \int_{\mathcal{L}_{\boldsymbol{\theta},\mathbf{t}}} f(x,y,z)\ dl, \quad (2)$$

where $\mathcal{L}_{\boldsymbol{\theta},\mathbf{t}}$ stands for the projection ray at angle $\boldsymbol{\theta} = (\theta, \varphi)$ and detector position $\mathbf{t} \in \mathbb{R}^2$ (see Figure 1). The parallel projection $\mathbf{p}_k$ of the image at a fixed angle $\boldsymbol{\theta} = \boldsymbol{\theta}_k$ ($k = 1, \ldots, m$) is a vector of size $T \times 1$, yielded by the set of projection values $p_{\boldsymbol{\theta}_k}(\mathbf{t})$ for T discrete values of $\mathbf{t}$, and the global projection vector $\mathbf{p} = [\mathbf{p}_1^t, \ldots, \mathbf{p}_m^t]^t$, corresponding to a set of m projection angles, is the result of the concatenation of all vectors $\mathbf{p}_k$.

0-7803-9134-9/05/$20.00 ©2005 IEEE

Let us discretize image $f(x, y, z)$ into a set of cubic voxels of density $\mathbf{f} = [f_1, \ldots, f_n]^t$. The X-ray operator rereads $\mathbf{p} = \mathbf{A}\mathbf{f}$, where $\mathbf{A}$ is a $M \times n$ projection matrix, $M = m\,T$ is the number of detector pixels, and n is the number of voxels.

Finally, the measurement model can be written as $\mathbf{d} = \mathbf{A}\mathbf{f}+\mathbf{n}$, where $\mathbf{d}$ stands for the X-ray data, $\mathbf{f}$ is the unknown image parameter vector, and the noise $\mathbf{n}$ accounts for both errors of modeling and measurement. For the sake of simplicity, $\mathbf{n}$ is assumed to be independent identically distributed (i.i.d.), white and Gaussian.

2.2. Monoresolution reconstruction

The monoresolution approach amounts to directly estimate the voxel densities $\widehat{\mathbf{f}}$ given the data $\mathbf{d}$. When the number or angles of projection are limited, this inverse problem is ill-posed due to the lack of data. We then use the regularization theory, and the Bayesian inference framework. We define the MAP estimate, as the minimizer of the compound energy:

$$\widehat{\mathbf{f}} \;=\; \underset{\mathbf{f}\in\mathbb{R}^n_+}{\arg\min}\, \big\{\mathcal{J}(\mathbf{f}) \;=\; \|\mathbf{d}-\mathbf{A}\mathbf{f}\|^2 + \lambda\mathcal{D}(\mathbf{f}) + \mu\mathcal{U}(\mathbf{f})\big\}. \quad (3)$$

Criterion $\mathcal{J}(\mathbf{f})$ is formed of a fidelity to data term, and two regularization terms, accounting for the homogeneity of voxel values, and the knowledge of the background value, respectively. The latter energies are defined by $\mathcal{D}(\mathbf{f}) = \sum_{i\sim j}\phi(f_i - f_j)$, where $\sim$ denotes the 6-neighborhood relationship between voxels, and $\mathcal{U}(\mathbf{f}) = \sum_{i=1}^n f_i$. Function ϕ is selected in order to favor piecewise homogeneous regions. We choose the Huber function, which is quadratic at the origin and linear at infinity:

$$\phi(t) \;=\; \begin{cases} t^2/T^2 & \text{if } |t| < T, \\ 2|t|/T - 1 & \text{otherwise.} \end{cases} \quad (4)$$

As criterion $\mathcal{J}(\mathbf{f})$ is convex on $\mathbb{R}^n_+$, and its gradient is Lipschitz, it has a unique local minimum which is equal to its global minimum. We then perform the optimization stage using a local gradient descent algorithm, projected on the set of constraints $\mathbb{R}^n_+$ [4].

3. SPARSE IMAGE RECONSTRUCTION

3.1. Sparse parameterization

In this section, we focus on a sparse modeling of the image, relying on the prior knowledge of the voxels that lay inside the background ($f_i = 0$). A decision rule for such background detection will be discussed in Section 3.2.

Let $\mathcal{V} \subset \{1, \ldots, n\}$ be the set of indices of the *active* voxels, *i.e.,* the voxels that do not lay in the background ($f_i > 0$ for all $i \in \mathcal{V}$), and v be their number. Setting aside the background voxels, the image parameterization reduces to vector $\mathbf{f}_v = \{f_i,\, i \in \mathcal{V}\} \in \mathbb{R}^v_+$ extracted from $\mathbf{f}$. Up to a reordering of the voxels, the complete image parameterization writes $\mathbf{f} = [\mathbf{f}_v^t, \mathbf{0}^t]^t$, where $\mathbf{0}$ is the null vector of length $n - v$ (background voxels). Criterion $\mathcal{J}(\mathbf{f})$ rereads:

$$\mathcal{J}(\mathbf{f}) \;=\; \mathcal{J}_{\mathcal{V}}(\mathbf{f}_v) \;\equiv\; \|\mathbf{d}-\mathbf{A}_v\mathbf{f}_v\|^2 + \lambda\,\mathcal{D}_{\mathcal{V}}(\mathbf{f}_v) + \mu\,\mathcal{U}_{\mathcal{V}}(\mathbf{f}_v), \quad (5)$$

where $\mathbf{A}_v$ is the new projection matrix, extracted from $\mathbf{A}$, of size $M \times v$. We also have $\mathcal{D}_{\mathcal{V}}(\mathbf{f}_v) = \mathcal{D}(\mathbf{f})$ and $\mathcal{U}_{\mathcal{V}}(\mathbf{f}_v) = \mathcal{U}(\mathbf{f}) = \sum_{i\,\in\mathcal{V}} f_i$. We choose to minimize $\mathcal{J}_{\mathcal{V}}(\mathbf{f}_v)$ over $\mathbb{R}^v_+$:

$$\widehat{\mathbf{f}}_v \;=\; \underset{\mathbf{f}_v\in\mathbb{R}^v_+}{\arg\min}\; \mathcal{J}_{\mathcal{V}}(\mathbf{f}_v), \quad (6)$$

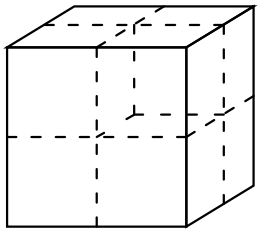

Figure 2: Multiresolution scheme. At each resolution level r, every voxel is divided into 8 identical sub-voxels of same density.

in order to approximate the minimizer $\widehat{\mathbf{f}}$ of $\mathcal{J}$ by $[\widehat{\mathbf{f}}_v^t, \mathbf{0}^t]^t$. In the same way as criterion (3), functional $\mathcal{J}_{\mathcal{V}}$ is convex, and its gradient is Lipschitz. Consequently, we can utilize a first order projected gradient algorithm to estimate $\widehat{\mathbf{f}}_v$.

We now have to define a criterion for deciding which voxels are active and which ones belong to the background. We answer to this question in the following, by using a multiresolution strategy.

3.2. Multiresolution scheme

We now design a multiresolution scheme, which yields successive sparse models for coarse to fine images, parameterized by $\mathbf{f}^{(1)} \in \mathbb{R}^{v_1}_+, \mathbf{f}^{(2)} \in \mathbb{R}^{v_2}_+, \ldots, \mathbf{f}^{(R)} \in \mathbb{R}^{v_R}_+$, where R is the number of resolution levels. Considering all active and background voxels put together, the R images are composed of $n,\, 8\,n,\, 8^2\,n, \ldots, 8^{R-1}\,n$ voxels, respectively. In other words, at a fixed level $r > 1$, each voxel of the image parameterized by $\mathbf{f}^{(r-1)}$ is divided into 8 sub-voxels (see Figure 2). Successive estimation of $\widehat{\mathbf{f}}^{(1)}, \widehat{\mathbf{f}}^{(2)}$, *etc.* consists in minimizing appropriate energies $\mathcal{J}_{\mathcal{V}_1}(\mathbf{f}^{(1)}),\ \mathcal{J}_{\mathcal{V}_2}(\mathbf{f}^{(2)})$, *etc.*, using the method presented above. In the following, we define a rule for recursive detection of the sets of active voxels $\mathcal{V}_r \subset \{1, \ldots, 8^{r-1}\,n\}$, and construction of respective energies $\mathcal{J}_{\mathcal{V}_r}(\mathbf{f}^{(r)})$.

We consider a fixed level $r > 1$, and assume the knowledge of $\mathcal{V}_{r-1}$ and $\widehat{\mathbf{f}}^{(r-1)} = \{\widehat{f}_k^{(r-1)},\ k \in \mathcal{V}_{r-1}\}$. For each voxel $i \in \big\{1, \ldots, 8^{r-1}\,n\big\}$, let $k \in \big\{1, \ldots, 8^{r-2}\,n\big\}$ be the father of i, that is the larger voxel which contains i.

- If $k \notin \mathcal{V}_{r-1}$, then i is set as a background voxel at level r.
- Else if $\widehat{f}_k^{(r-1)} = 0$, then i is set as a background voxel at level r.
- Else, i is set active at level r, and $\widehat{f}_i^{(r)}(0) = \widehat{f}_k^{(r-1)}$.

This simple rule allows us to define $\mathcal{V}_r$ and the initial solution $\widehat{\mathbf{f}}^{(r)}(0)$ to the estimation of $\widehat{\mathbf{f}}^{(r)}$, from solution $\widehat{\mathbf{f}}^{(r-1)}$. Both vectors $\widehat{\mathbf{f}}^{(r-1)} \in \mathbb{R}^{v_{r-1}}_+$ and $\widehat{\mathbf{f}}^{(r)}(0) \in \mathbb{R}^{v_r}_+$ describe the same image at resolution levels $r-1$ and r.

At the first level $r = 1$, we set $\mathcal{V}_1 = \{1, \ldots, n\}$ (all voxels are active), where n is an arbitrary number of voxels, and we define $\widehat{\mathbf{f}}^{(1)}(0)$ as the n voxel backprojection of the data $\mathbf{d}$.

Table 1 summarizes the successive parameterization of the image at levels $1, \ldots, R$, and the estimation of solutions $\widehat{\mathbf{f}}^{(r)}$ from vectors $\widehat{\mathbf{f}}^{(r)}(0)$. At resolution level r, the cost function $\mathcal{J}_{\mathcal{V}_r}(\mathbf{f}^{(r)})$ is formulated using (5), and minimized by the local projected gradient descent algorithm. Hyperparameters are set to $\lambda_r = \lambda_{r-1}/4$, $\mu_r = \mu_{r-1}/8$, and $T_r = T_{r-1}$, in order to guarantee that $\mathcal{J}_{\mathcal{V}_r}\big(\widehat{\mathbf{f}}^{(r)}(0)\big) = \mathcal{J}_{\mathcal{V}_{r-1}}\big(\widehat{\mathbf{f}}^{(r-1)}\big)$. As each minimization of $\mathcal{J}_{\mathcal{V}_r}$ involves a descent algorithm, we have, for all $r > 1$, $\mathcal{J}_{\mathcal{V}_r}\big(\widehat{\mathbf{f}}^{(r)}\big) \leqslant \mathcal{J}_{\mathcal{V}_{r-1}}\big(\widehat{\mathbf{f}}^{(r-1)}\big)$. Therefore, the multiresolution scheme yields a

Perform backprojection $\widehat{\mathbf{f}}^{(1)}(0)$ of the data $\mathbf{d}$ on a grid on n voxels. Set $\mathcal{V}_1 = \{1, \ldots, n\}$ (all voxels are active).
For $r = 1, \ldots, R$,
  Estimate $\widehat{\mathbf{f}}^{(r)} = \arg\min_{\mathbb{R}_+^{v_r}} \mathcal{J}_{\mathcal{V}_r}(\mathbf{f}^{(r)})$ from $\widehat{\mathbf{f}}^{(r)}(0)$.
  If $r < R$,
    Define $\mathcal{V}_{r+1}$, by discretizing the non null voxels of $\widehat{\mathbf{f}}^{(r)}$.
    Compute $\widehat{\mathbf{f}}^{(r+1)}(0)$ from solution $\widehat{\mathbf{f}}^{(r)}$.

Table 1: Multiresolution algorithm for voxel reconstruction.

sequence of solutions that correspond to decreases of the values of criterion $\mathcal{J}(\mathbf{f})$, where $\mathbf{f} \in \mathbb{R}_+^{8^{R-1}\,n}$ here stands for the complete image parameterization at the finest resolution level.

4. SIMULATION RESULTS

We illustrate and compare both monoresolution and multiresolution methods on a set of synthetic data. This simulation involves $m = 9$ noisy projections of a binary object, plotted on Figure 3. The angles of projection, specified on Table 2, are limited, as all angles φ_k are lower than $\pi/4$. The signal to noise ratio, defined as SNR $= 10 \log v_p/v_n$, where v_p and v_n are the empirical variances of the projection signal and the noise, is set to 10 dB.

Monoresolution reconstruction involves an image composed of $n = 64^3$ voxels. The estimate $\widehat{\mathbf{f}}$, plotted on Figure 4, is obtained with the projected gradient descent algorithm (15 iterations), and with the data backprojection as initial solution $\widehat{\mathbf{f}}(0)$. Hyperparameters μ and T are set to $\mu = \lambda/3$, and $T = 1/5$, and λ is selected such that $\|\mathbf{d} - \mathbf{A}\widehat{\mathbf{f}}(0)\|^2 = 0.9\,\mathcal{J}\big(\widehat{\mathbf{f}}(0)\big)$. Although the quality of reconstruction is acceptable, the reconstructed image contains many close to zero, however non zero valued voxels. The global time of reconstruction amounts to 4397 seconds (CPU time), and would become enormous for larger values of n (see Figure 6 (b)).

Multiresolution results are obtained using three levels of resolution, corresponding to $n = 16^3$, 32^3 and 64^3 active and background voxels (see Figure 5). The number of active voxels is $v_1 = 16^3 = 4096$, $v_2 = 5512$, and $v_3 = 32656$ for each resolution level; v_2 and v_3 represent 16.8% and 12.5% of the total number of voxels n. At level $r = 1$, hyperparameters are set as described in the previous paragraph. Their selection is automatic for $r = 2$ and 3, as $\lambda_r = \lambda_{r-1}/4$, $\mu_r = \mu_{r-1}/8$, and $T_r = 1/5$. The global time of reconstruction amounts to 852 seconds. As r increases, the number v_r of active parameters remains low, and the initial solution $\widehat{\mathbf{f}}^{(r)}(0)$ is close to the local minimizer $\widehat{\mathbf{f}}^{(r)}$ of $\mathcal{J}_{\mathcal{V}_r}$. Consequently, the number of iterations needed for convergence of the gradient descent algorithm is very limited, and the reduction of the reconstruction time is drastic (see Figure 6).

k	1	2	3	4	5	6	7	8	9
θ_k	$\pi/4$	$3\pi/4$	$-\pi/4$	$5\pi/4$	$-\pi/2$	$\pi/2$	0	π	0
φ_k	φ	φ	φ	φ	$\pi/4$	$\pi/4$	$\pi/4$	$\pi/4$	0

Table 2: Values of projection angles. Angle $\varphi = \arccos(\sqrt{3}/3)$.

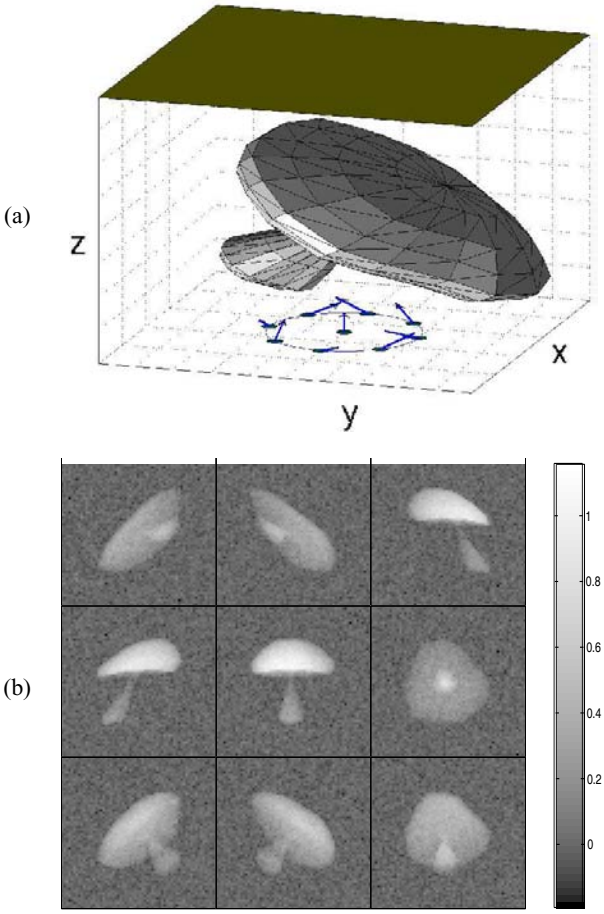

Figure 3: Simulation data: 9 limited angle projections of a mushroom-shape binary object. SNR is set to 10 dB. (a) Binary synthetic object and projection geometry, (b) The 9 projection images, of size 64×64, put together in a 3×3 "matrix". The first row corresponds to the first three projection images, *etc.*

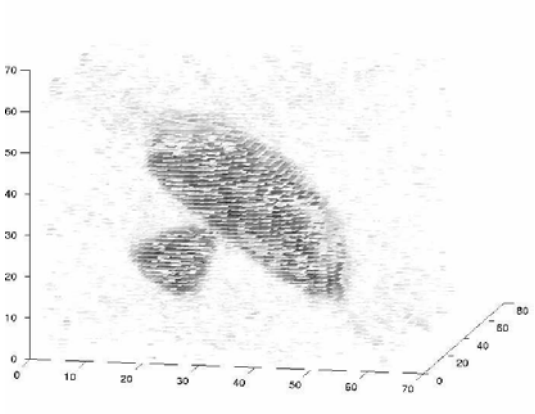

Figure 4: Monoresolution result $\widehat{\mathbf{f}}$, formed of $n = 64^3$ voxels. Estimation of $\widehat{\mathbf{f}}$ is done by performing 15 iterations of the projected gradient algorithm, and with the backprojection of the data as initial solution. The time of reconstruction is 4397 seconds (CPU).

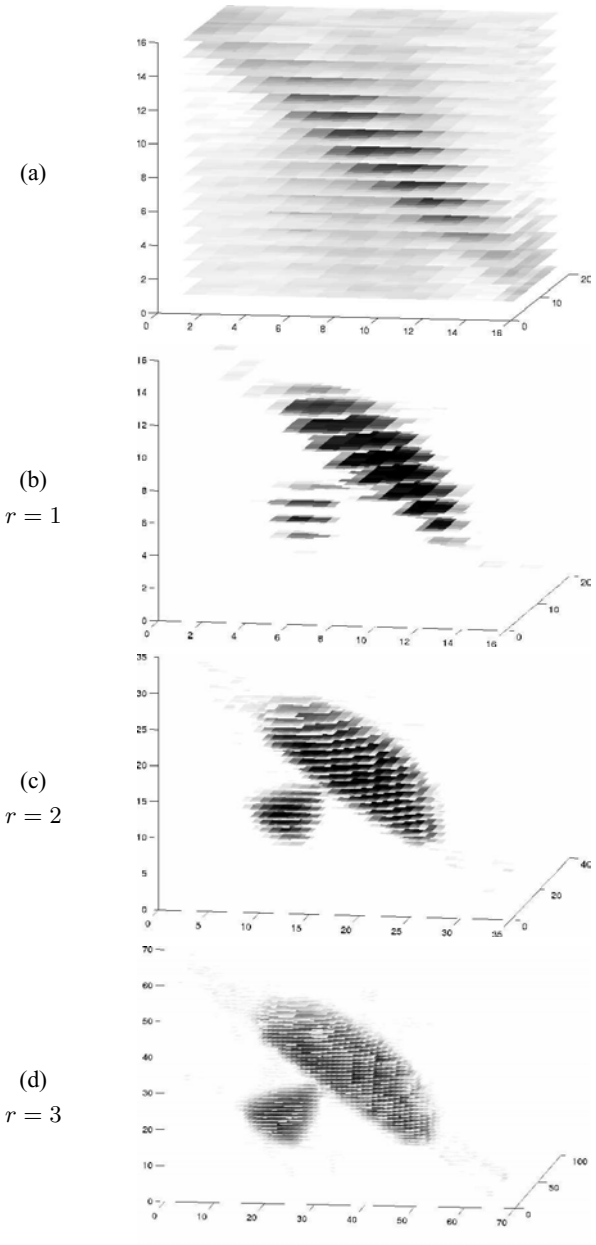

Figure 5: Multiresolution reconstructions on three levels. (a) Data backprojection $\widehat{\mathbf{f}}^{(1)}(0)$, computed on a grid of $n = 4096$ voxels. (b,c,d) Reconstructions $\widehat{\mathbf{f}}^{(1)}, \widehat{\mathbf{f}}^{(2)}, \widehat{\mathbf{f}}^{(3)}$, corresponding to p^3 (active and background) voxels, where $p = 16, 32$, and 64.

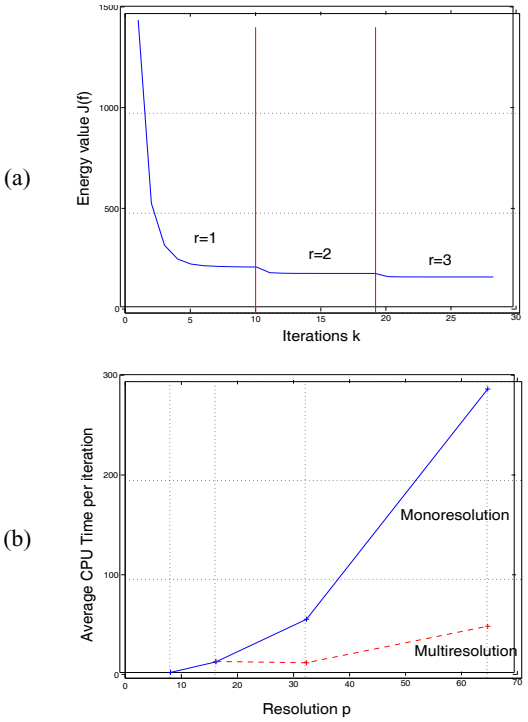

Figure 6: Quality of reconstructions. (a) Plot of criterion values $\mathcal{J}(\mathbf{f})$ for the multiresolution reconstruction, on three levels $r = 1, 2, 3$. At each level, the projected gradient descent algorithm is utilized, with 10 iterations ($r = 1$) and 8 iterations ($r = 2, 3$). (b) Plot of the average CPU time per iteration of the gradient algorithm, in mono and multiresolution cases. $p = \sqrt[3]{n} = 16, 32, 64$ is the image resolution, CPU time is expressed in seconds.

5. CONCLUSIONS

We have presented an original method for reducing the number of parameters involved in 3D image reconstruction from X-ray data. Our method relies on a sparse description of the image, and a multiresolution technique, in order to reconstruct a sequence of coarse to fine sparse images. The multiresolution scheme affords an automatic detection of the background voxels, as well as a good initial estimate of the active voxels at each resolution level. Therefore, MAP estimation of the active voxel densities necessitates a very low number of iterations of the local optimization algorithm, and a reduced amount of CPU time. The method can be trivially extended to the case where the background is not uniform, but known. Such images can be encountered in NDE applications, where the background corresponds to known materials, and one wants to reconstruct air faults. We can expect that the multiresolution strategy will bring drastic amelioration, since faults are generally of small size, relative to the entire structure.

6. REFERENCES

[1] S. Gautier, J. Idier, A. Mohammad-Djafari, and B. Lavayssière, "X-ray and ultrasound data fusion", in *Proc. IEEE ICIP*, Chicago, IL, Oct. 1998, pp. 366–369.

[2] A. B. Frakt, W. C. Karl, and A. S. Willsky, "A multiscale hypothesis testing approach to anomaly detection and localization from noisy tomographic data", *IEEE Trans. Image Processing*, vol. 7, no. 6, pp. 825–837, June 1998.

[3] J. C. Ye, C. A. Bouman, K. J. Webb, and R. P. Millane, "Nonlinear multigrid algorithms for Bayesian optical diffusion tomography", *IEEE Trans. Image Processing*, vol. 10, no. 6, pp. 909–922, June 2001.

[4] D. P. Bertsekas, *Nonlinear programming*, Athena Scientific, Belmont, MA, 1995.