1

# 3D Reconstruction in X-ray Tomography: Approach Example for Clinical Data Processing

# 1.1. Introduction

Works presented in this chapter stem from three-dimensional (3D) reconstruction problems, in X-ray computed tomography (XRCT), within a clinical framework. More specifically, the practical objective was to use XRCT to detect and quantify possible restenosis occurring in some patients after the insertion of a stent. The quality of reconstructions achieved by clinical tomographs being insufficient for this purpose, the concern was thus to develop a method capable of rebuilding small structures in the presence of metal objects, in a 3D context, with a precision higher than that of tomographs available in hospitals; in addition, this method was supposed to work with computers commonly available in most research laboratories (such as personal computers (PCs), without any particular architecture or processor hardware).

The development of a solution clearly falls within the framework of conventional inverse problem solving. However, it is essential to take into account the characteristics of 3D XRCT, and notably the very large volume of data to be processed, the geometric complexity of the collection process of raw data, and practical barriers to access these data. In order to achieve the objective stated above, difficulties are twofold: (1) at the methodological level, the development of an inversion method adapted to the intrinsic characteristics of the problem to be addressed and, (2) at the implementation level, accounting for the practical obstacles mentioned above as well as for constraints on the processing time inherent in any

Chapter written by Yves GOUSSARD.

clinical application. In the following, we present the retained approach based on the analysis of the main factors likely to improve the quality of reconstructions while satisfying the practical constraints which we must face; this brings us to putting the methodological aspects into perspective, with regard to practical questions, in light of the applied objective of these works.

### **1.2.** Problem statement

Although image reconstruction methods used in the first tomographs were of the analytical type [AMB 73, HOU 73], the advantages of approaches based on estimation [HER 71, HER 73, HER 76b], then the ill-posed nature of tomographic reconstruction problems [HER 76a, HER 79] were recognized very early on. Over the past 35 years, many academic studies focusing on tomographic reconstruction have been carried out in the context of solving inverse problems. Generally, the emphasis is on the main three elements of this type of approach, that is to say, modeling of the data formation process, choice of the estimator, and development of techniques that enable the practical computation of the estimate [DEM 89]. These works have been partly customized according to various imaging modalities (for example, transmission [HER 76a], emission [LEV 87], diffraction [BER 97] tomography, and more recently optical and/or multiphysics tomography (see [BOA 04] for a partial synthesis)) which present largely variable degrees of if estimation conditions are often very unfavorable in diffraction difficulty: tomography (eddy current tomography [TRI 10], seismic imaging [VAU 11]) due to the strong non-linearity of underlying physical phenomena, to the importance of attenuation phenomena and to the small number of observations with respect to the number of unknowns, the inversion conditions are generally better in emission tomography (SPECT for example) and can be qualified as relatively favorable in XRCT. This explains why, in this area, reconstruction methods known as "naive" provide results that have been used in clinical practice for several decades. Thereafter, we present the elements likely to have a significant impact on the performance of an XRCT inversion method.

#### **1.2.1.** Data formation models

In XRCT, all data formation models are based on the Beer–Lambert law, which describes the attenuation of an X-ray beam through a medium whose spatial distribution of the attenuation coefficients is referred to by  $\mu$ . It takes the form:

$$N: \mathcal{P}\left\{n_0 \exp\left\{-\int_D \mu(s) \, \mathrm{d}s\right\}\right\}$$
[1.1]

where N is the random variable <sup>1</sup> that represents the number of photons arriving at the detector,  $n_0$  the number of photons emitted by the source and D the path (straight line) of photons between the source and the detector.  $\mathcal{P} \{\ell\}$  refers to the Poisson distribution with parameter  $\ell$ . By discretizing the distribution of the attenuation coefficients of the medium, the integral involved in [1.1] becomes:

$$\int_D \mu(s) \, \mathrm{d}s = \boldsymbol{a}_D^{\mathrm{t}} \boldsymbol{\mu} \tag{1.2}$$

where  $\mu$  is the vector in which all samples of the attenuation coefficients of the medium are concatenated, and where  $a_D$  is a vector representing the contribution of each sample of  $\mu$  to the integral. By concatenating the variables N corresponding to each source-detector position into vector N, and by concatenating in an analogous manner row vectors  $a_D^t$  in matrix A, one obtains:

$$\boldsymbol{N}: \mathcal{P}\left\{n_0 e^{-\boldsymbol{A}\boldsymbol{\mu}}\right\}$$
[1.3]

where the Poisson distribution should be understood component-wise, all components being independent from each other. A represents the projection matrix of the tomographic system which is sparse, usually very large and structured in a way that reflects the geometry of the data collection process.

In most methods, reconstruction is carried out not from the photon counts, but from the auxiliary quantity  $\mathbf{Y} \triangleq -\ln \mathbf{N}/n_0$ . By performing a Taylor-series expansion of the neg-log-likelihood which derives from [1.3], Sauer and Bouman [SAU 93, Appendix A] have shown that it assumes the following approximate form:

$$J_{\mathbf{V}} \propto \frac{n_0}{2} \left\| \boldsymbol{y} - \boldsymbol{A} \boldsymbol{\mu} \right\|_{\boldsymbol{\Sigma}}^2, \quad \boldsymbol{\Sigma} \triangleq \operatorname{Diag} \left\{ e^{-\boldsymbol{y}} \right\}$$
[1.4]

In practice, this amounts to assuming that:

$$\boldsymbol{Y} \approx \mathcal{N} \left\{ \boldsymbol{A} \boldsymbol{\mu} , 1/n_0 \operatorname{Diag} \left\{ e^{\boldsymbol{y}} \right\} \right\}$$
[1.5]

where  $\mathcal{N} \{m, Q\}$  denotes the normal distribution with mean m and covariance matrix Q. This approximation allows the estimation of  $\mu$  to be carried out in a linear Gaussian framework, the problem remaining generally very large.

The models presented above are valid when X-ray sources and detectors are punctual and when the radiation emitted by the source is monoenergetic. However,

<sup>1.</sup> In this chapter, any random quantity is designated by an uppercase character and its realization by the corresponding lowercase character.

these two assumptions are relatively rough approximations of reality, and their impact on the quality of reconstructed images can be significant in practice.

On the one hand, considering sources and detectors are punchual, while they have a non-zero surface, oscillations, which are all the larger as the discretization of the medium is fine, are produced in the reconstructed images . In addition, these artifacts are highly structured as they reflect the paths of radiation in the medium. In an attempt to limit the magnitude of these phenomena, several techniques aimed at changing the structure of matrix A in order to account for the thickness or the volume of the ray beams have been proposed [DEM 04, ZHU 94]. However, it can be observed that these techniques have a strong empirical character, insofar that they all overlook the fact that the calculation of the parameter of the Poisson distribution of photon counts requires the summation of exponential functions (see equation [1.1] or [1.3]), and cannot be carried out directly on the linear Gaussian model [1.5]. Nevertheless, in practice, these approaches limit the appearance of artifacts related to the thin ray model.

On the other hand, the hypothesis of monoenergetic X-ray source is never fulfilled in practice, since most of the sources used in XRCT have a spectrum that ranges from 0 to approximately 120 keV. However, the effect of this hypothesis only appears when variations of  $\mu$  with respect to energy are significant. In moderately attenuating media such as water and soft tissue, these variations are small, and the hypothesis of a monoenergetic radiation source only causes minimal degradations. On the other hand, in more strongly attenuating media (cortical bone, metal, vascular calcifications), variations are more significant, and the monoenergetic assumption results in various artifacts (star-shaped motifs, artificial thickening of strongly attenuating structures) that can make it difficult, if not impossible, to interpret the results. Unfortunately, there is a rather scarce number of works aimed at accounting for the polyenergetic nature of radiation in reconstruction methods in an explicit and general manner. However, we should emphasize the contribution of [DEM 01], based on the Alvarez-Macovski decomposition of attenuation coefficients [ALV 76] and on an empirical parameterization of the variations of these coefficients with respect to energy. We will refer in more detail to this model in the next section. It should be noted that, as for accounting for the non-zero surface of detectors, determining the parameter of the Poisson distribution of polyenergetic radiation requires the summation of exponential functions and involves strong nonlinearities.

Regardless of the type of data formation model used, the product by A (projection operation) dominates the computation of the prediction of the observations for a given medium. Most inversion techniques (direct or iterative) also require the product by  $A^{t}$ . Due to the large size of the problem, the choice of the representation of A (implicit or explicit, accessible by rows or by columns), and the effectiveness of the computation of the left- and right-products by A have a critical impact on the effectiveness of reconstruction methods. It should be noted that these issues are rarely addressed in the open literature.

## 1.2.2. Estimators

As in most of the other modalities, the vast majority of XRCT reconstruction methods based on an inversion approach rely on penalized maximum likelihood estimators. The likelihood portion derives from the distribution of photon counts [1.3] or from the Gaussian approximation [1.5] with respect to variable Y. In either case, the resulting neg-log-likelihood is convex under the monoenergetic hypothesis and is well-suited for minimization by descent methods. Under the polyenergetic hypothesis, the neg-log-likelihood takes a more complex form, and loses the convexity property due to the strong nonlinearity of the parameterizations used. These elements are a likely cause of the scarcity of works based on polyenergetic models.

With regard to the penalty portion, or *a priori* modeling of the medium to reconstruct, a broad consensus can be observed in employing Markovian representations with convex potential on first-differences [BOU 93, DEL 98, FES 94]. This is reflected in the neg-log-likelihood by the addition of a convex additive penalty term, quadratic or not. Surprisingly, there is to our knowledge no real study comparing, in XRCT and under realistic conditions, the effect of quadratic (Gaussian Markov random fields) and non-quadratic (edge-preserving) penalty terms. In practice, the latter are usually used because they better reflect the global properties of media to be reconstructed at a very low marginal cost.

## 1.2.3. Algorithms

Reconstruction algorithms refer here to numerical optimization procedures that allow us to minimize the objective functions (usually penalized likelihood) referred to in the preceding paragraph. As previously mentioned, these objective functions have been developed under monoenergetic assumptions, and the same assumptions are in force for algorithms used in XRCT. However, it should be noted that a number of optimization procedures developed for other modalities (for example, emission computed tomography) can easily be transposed to XRCT provided that the data formation model is linear.

Since the introduction of the first "algebraic" reconstruction methods [HER 71], development of optimization techniques aimed specifically at tomographic reconstruction has been the subject of numerous works. The strength of this activity, and the fact that generic nonlinear optimization methods have been widely ignored by the community, can be attributed to the practical difficulties in solving reconstruction problems (large size, need to account for bound constraints) and the general perception that generic optimization methods were unable to cope with these problems in a satisfactory manner. Most algorithms specific to emission or to transmission tomography fall into the framework of iterative descent direction methods, and require

the computation of the gradient of the objective function (or of one of its partial derivatives) at each iteration. It is not possible to detail here the various approaches that lead to algebraic structures suited for such and such specific representation of matrix A. Among the most significant works, ICD algorithms [SAU 93], as well as approaches based on ordered subsets (OS) that help to significantly reduce the computation time, can be mentioned. However, these works appear to present two major flaws:

- studies that compare the different approaches are rare and extremely narrow in scope. If carrying out such studies in a rigorous manner presents significant difficulties related to the possible differences of implementation, to the possibility of parallelization, etc., the adoption of a common framework would nevertheless make it easier to identify major trends. Today, generic optimization methods no longer present the same limitations as in the past and lend themselves relatively well to solving reconstruction problems, which makes the lack of such comparative studies even more unfortunate;

- the issues of convergence control of reconstruction methods are scarcely or not addressed, which gives the results a pronounced empirical character and makes comparison attempts between methods even more difficult. In particular, OS-based methods are not convergent, which leaves unanswered questions about the choice of a stopping criterion and its link with the quality of the reconstructed images.

For these reasons, the choice of an XRCT optimization algorithm still seems to be lacking of rigor and to present a significant empirical character.

Despite the abundance and quality of academic works, their penetration in XRCT clinical applications appears to be minimal. As an example, the recent works dedicated to applications focus mostly on 3D computerized axial tomography (systems such as C-Arm and O-Arm) [KAC 11], as opposed to helical tomographs available in most hospitals. This situation is partly explained by the fact that, for nearly 30 years, the four major manufacturers of clinical tomographs have focused almost all of their efforts on hardware improvement and on the timeliness of obtaining results rather than on the reconstruction methodology. The result is that, today, measurements are of satisfactory quality and the conditioning of the reconstruction problem is acceptable, which limits the need to resort to reconstruction techniques taking full advantage of the measurements. Moreover, the large volume of data to process, the geometric complexity of data collection and processing time constraints make academic works difficult to apply, and their real contribution sometimes questionable. This is the reason why iterative reconstruction methods are just beginning to appear in clinical tomographs, in a very elementary form.

The achievement of the objectives formulated in the introduction requires a range of choices related to the methodology as well as to the practical aspects of reconstruction in XRCT; the coherence of these choices is essential for obtaining a satisfactory performance. The situation that we have briefly outlined reveals that the tools to guide these choices are relatively scarce, and limited in scope. In what follows, we present how we have made these choices as well as the resulting method.

# 1.3. Method

Recall that the objective is to develop a 3D reconstruction XRCT method, capable of processing data generated by clinical tomographs and of creating reconstructed images of better quality than these latter. In addition, the method should work with current equipment such as PCs, and consistency in the choice of its various components appears as essential to achieve these goals. In this section, we detail and discuss these choices.

## 1.3.1. Data formation models

One of the critical elements in the choice of a data formation model is accounting, or not, for the polyenergetic nature of the X-ray source and the dependence of  $\mu$  with regard to energy. However, as mentioned earlier, studies dedicated to polyenergetic models are rare, and even more those that allow the assessment of the respective contribution of monoenergetic and polyenergetic models within an inversion framework. Since the studied medium involved strongly attenuating objects (the *stent* metal parts), it appeared necessary to develop methods based on both models, and to compare their effectiveness for the targeted application. Another important element is accounting for the non-zero surface of the source and of the detectors: artifacts linked to a thin ray assumption appear during the reconstruction of high resolution images, and that is precisely the situation in which we find ourselves when trying to accurately image small-sized vascular structures.

Even before developing the models in question, it is necessary to select a representation of the projection matrix A, the important points being the explicit (prior computation) or implicit (on-the-fly computation) character of the operator, the possibility of easily accessing rows or columns and, finally, the development of an internal representation adapted to the retained formulation. We have opted for an explicit representation of the projection operator on a single rotation, since such a formulation leads to a lower amount of computation than implicit representations, at the price however of significant increase in memory space. Limiting the operator to a single rotation introduces an additional constraint (integer number of slices of the object per rotation) that is scarcely penalizing in practice. In order to limit the memory space required for the storage of the operator, we have developed a coding system for A that takes advantage of its sparse and structured character, that allows quick access to the rows of the matrix and that lends itself to the efficient

computation of projection and backprojection operations [GEN 08, GEN 11]. However, for the usual configurations of clinical tomographs, the storage of A requires several GB. An approach to circumventing this difficulty is detailed in section 1.3.3.

We now develop a formulation that allows both the thickness of rays and the polyenergetic character of the source to be considered. To this end, it can be observed that accounting for any of these phenomena amounts to performing a summation on the parameter of the Poisson distribution of photon counts given in [1.3]. More specifically, by using the independence of attenuation phenomena of different photons, and the basic properties of the Poisson distribution, [1.3] becomes after discretization:

$$\boldsymbol{N}: \mathcal{P}\left\{n_0 \sum_{k=1}^{K} \alpha_k e^{-(\boldsymbol{A}\boldsymbol{\mu})_k}\right\}$$
[1.6]

In order to account for the thickness of rays, and under the hypothesis that only detectors have a nonzero surface (the geometry of the source is rarely available), it is considered that infinitely thin rays reach each detector in K distinct positions, which may be determined by regular sampling or random selection. In the above equation, it then gives:

$$\alpha_k = 1/K$$
 and  $(A\mu)_k = A_k\mu$  [1.7]

and  $A_k$  represents the projection matrix under the hypothesis of thin rays obtained for the k-th configuration of the position of rays on the detectors.

Developing a polyenergetic model requires taking into account the dependence of attenuation coefficients with respect to energy. Discretizing the spectrum of the source on K levels, and denoting by  $\mu_k$ , the attenuation coefficients of the object at the k-th energy level yields:

 $\alpha_k$ : discretized spectrum of the source

 $(\boldsymbol{A}\boldsymbol{\mu})_k = \boldsymbol{A}\boldsymbol{\mu}_k$ 

The fact that the object to reconstruct is no longer represented by  $\mu$ , but by the K attenuation coefficients maps  $\{\mu_k \ 1 \le k \le K\}$  leads both to a very large underdetermination of the reconstruction problem, and to difficulties in the interpretation of the results due to redundancy of the information present in the different maps. The Alvarez–Macovski decomposition enables  $\mu_k$  to be expressed as

a function of the attenuation due to the photoelectric effect on the one hand, and of that due to the Compton effect on the other hand. More specifically, it gives:

$$\mu_k = \Phi(k)\phi + \Theta(k)\theta$$

where  $\Phi(k)$  and  $\Theta(k)$  are known deterministic functions, and where  $\phi$  and  $\theta$  respectively denote the photoelectric and Compton coefficients [DEM 01]. This decomposition makes it possible to represent the object by the two maps  $\phi$  and  $\theta$  rather than by the K maps  $\mu_k$ . In order to end up with a single map, an empirical parameterization of  $\theta$  and  $\phi$  is introduced on the basis of a single quantity, in general  $\mu_{70}$  (attenuation at 70 keV), this parameterization being determined from the properties of a set of materials likely to be present in the object to be reconstructed [DEM 01]. Finally, we obtain:

$$\alpha_k$$
: discretized spectrum of the source [1.8a]

$$(\boldsymbol{A}\boldsymbol{\mu})_{k} = \boldsymbol{A}\left(\Phi(k)\boldsymbol{\phi}(\boldsymbol{\mu}_{70}) + \Theta(k)\boldsymbol{\theta}(\boldsymbol{\mu}_{70})\right)$$
[1.8b]

and it can be observed that this formulation is equivalent to performing a partial decoupling of the dependence of the attenuation coefficients with respect to space, on the one hand (variable  $\mu_{70}$ ), and with respect to energy, on the other hand (functions  $\Phi$  and  $\Theta$ ).

The previous developments helped to establish a common formulation to take into account the thickness of rays and the polyenergetic character of the X-ray source. The next step consists of using this formulation to develop an estimator of the attenuation coefficients. The most immediate approach is to use the Poisson neg-log-likelihood of the vector of observed photons counts n, which is directly derived from [1.6], [1.7] and [1.8]. Unfortunately, the expressions of this neg-log-likelihood and of its gradient with respect to the quantity of interest are complex, and involve transcendental functions whose practical computation is incompatible with a fast and effective implementation. To overcome this difficulty, the Poisson neg-log-likelihood is approximated by its second Taylor expansion around  $y \triangleq -\ln n/n_0$ , in the same spirit as [SAU 93, Appendix A]. The  $J_V$  criterion that derives therefrom takes the form:

$$J_{\mathrm{V}} = \frac{n_0}{2} \left\| \boldsymbol{y} - \sum_{k=1}^{K} \alpha_k (\boldsymbol{A}\boldsymbol{\mu})_k \right\|_{\boldsymbol{\Sigma}}^2$$
[1.9]

where  $\Sigma$  has been defined in [1.4].

The above expression deserves interpretation. When the objective is to account for the thickness of rays, [1.9] and [1.7] lead to:

$$J_{\mathrm{V}} = \frac{n_0}{2} \left\| \boldsymbol{y} - \left( \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{A}_k \right) \boldsymbol{\mu} \right\|_{\boldsymbol{\Sigma}}^2$$
[1.10]

Comparing [1.10] to [1.5], it can be observed that accounting for the thickness of rays amounts to using as a projection matrix the mean of matrices  $A_k$ . This type of procedure is similar to some of the empirical approaches mentioned in section 1.2.1 (methods known as *ray-driven*) and therefore constitute a form of justification. The link with other empirical approaches (including the methods known as *pixel-driven*) and *distance-driven*) seems more far-fetched. Our works will be directly based on [1.10].

In order to account for the polyenergetic character of the X-ray source, [1.9] and [1.8] give:

$$J_{\mathrm{V}} = \frac{n_0}{2} \left\| \boldsymbol{y} - \left( \sum_{k=1}^{K} \alpha_k \Phi_k \right) \boldsymbol{A} \boldsymbol{\phi}(\boldsymbol{\mu}_{70}) - \left( \sum_{k=1}^{K} \alpha_k \Theta_k \right) \boldsymbol{A} \boldsymbol{\theta}(\boldsymbol{\mu}_{70}) \right\|_{\boldsymbol{\Sigma}}^2$$
[1.11]

It can be observed that  $J_V$  has a much simpler structure than the corresponding Poisson neg-log-likelihood, and that the practical evaluation of its value and its gradient is relatively easy to implement: these operations require two projections and two backprojections, instead of one projection and one backprojection in the monoenergetic case. In general,  $\phi(\mu_{70})$  and  $\theta(\mu_{70})$  have a nonlinear behavior, which may complicate the computation of the gradient and degrade the numerical conditioning of the problem. However, we have observed that, in practice, these difficulties are not such as to call in question the advantages of the approach.

It should finally be mentioned that the approximation [1.9] makes it possible to easily and simultaneously account for the polyenergetic character of the X-ray source and the thickness of rays. As mentioned above, we will develop and compare methods based on monoenergetic and polyenergetic models ; in either case, ray thickness will be accounted for in accordance with [1.10].

## 1.3.2. Estimator

Irrespective of the type of data formation model retained, a maximum *a posteriori* (MAP) estimator is used due to its interesting trade-off between simplicity and adaptability to problems of variable difficulty. For the likelihood portion (or data fit), the quadratic approximation of the Poisson neg-log-likelihood is used, whose general

form is given in [1.9]. For the penalty portion, we use an additive, convex "edge-preserving" term of the form:

$$J_{\rm P} = \sum_{m=1}^{M} \varphi \left[ \boldsymbol{D}_m \boldsymbol{\mu} \right]$$
[1.12]

where  $\varphi$  is a scalar  $L_2L_1$  potential, where  $D_m$  refers to a – possibly weighted – matrix of first differences in a given direction m (horizontal, vertical, diagonal transverse, etc.) and where, for any vector  $u = \{u_n \ 1 \le n \le N\}$ , the notation  $\varphi[u]$  denotes the sum  $\sum_{n=1}^{N} \varphi(u_n)$ . Such a penalty term may be interpreted as an *a priori* Markov distribution of the attenuation coefficients field. With regard to the precise choice of the function  $\varphi$ , the following hyperbolic form is adopted for reasons of computational simplicity:

$$\varphi(u) = \sqrt{\delta^2 + u^2} \tag{1.13}$$

where  $\delta$  denotes a scaling hyperparameter. Under these hypotheses, performing the reconstruction amounts to minimizing the criterion  $J \triangleq J_V + \lambda J_P$  with respect to the variable of interest, where  $\lambda$  denotes the regularization parameter. In the case of a monoenergetic model, the variable of interest is  $\mu$  and J is convex, which lends itself to iterative minimization with a descent directions method. In the case of a polyenergetic model, the variable of interest is  $\mu_{70}$ , and the convexity of J cannot be guaranteed. However, we will make the assumption that nonlinearities in the model are sufficiently moderate to allow the use of the same minimization techniques as in the monoenergetic case.

### 1.3.3. Minimization method

#### 1.3.3.1. Algorithm selection

The minimization of J is an optimization problem of very large size, possibly nonlinear, convex or nearly convex, and subject to boundary constraints to ensure the positivity of the quantity of interest to reconstruct. Among the techniques capable of addressing such a problem, we restrict our choice to gradient-based, descent directions methods, due to their interesting trade-off between complexity and numerical effectiveness for the size of the problem under consideration. Within this type of techniques, the possibilities are multiple, both among generic methods (derived from numerical analysis and from optimization) and among algorithms specifically developed for tomography. The representation that we have adopted for the projection operator compels us to eliminate the methods requiring an easy access to the columns of A, such as the ICD method [SAU 93]. Even with these restrictions, candidate algorithms are still numerous, and the literature offers few objective elements capable of providing guidance in the choice of a particular method. That is why we have conducted a comparative study of limited scope about the effectiveness of several algorithms suited to our problem.

Developing a rigorous comparison methodology of iterative optimization algorithms for a given application is a non-trivial task. As a first step, it is necessary to specify an implementation framework adapted to the targeted application. Here, the focus is on sequentially programmed methods (or at least without any particular concern for parallelization) running on standard hardware such as PCs. This allows for the definition of performance metrics such as the total run-time, the number of projections and backprojections, etc. In this context, the objective is to achieve a given "image quality", relatively to the subsequent use of the results. This requires us to define image quality metrics such as resolution, ability to distinguish low contrasts, etc., and then to find a stopping rule, usable with all algorithms involved in the comparison, which ensures that the desired image quality is reached. It is essential that this stopping rule does not reflect the behavior of the algorithm, but the intrinsic properties of the solution. In other words, the stopping rule should depend on criterion J regardless of the algorithm used, which excludes some criteria commonly used such as thresholds on the norm of the solution updates or on the variation of the criterion between successive iterations. Clearly, a stopping rule related to optimality conditions satisfies the requirements of independence with respect to the algorithm, and a threshold based on the norm of the projected gradient is a candidate of choice because of its availability and computational simplicity. It should be noted that the use of this type of stopping rule eliminates non-convergent algorithms, and in particular ordered subsets type methods yet very widely used in tomographic reconstruction. Furthermore, comparison results should be independent of specific instances of the observations, which requires us to repeat the same experiment several times in order to perform a statistical comparison of the performance of the different methods.

In order to satisfy these requirements, we proposed a comparison methodology [HAM 11] based on simulated data derived from phantoms containing random elements that can be used to compute quality metrics of reconstructed images. The simulated data set allows for a statistical comparison of performance measures for each algorithm with a common stopping rule. In addition, a statistical comparison of quality metrics of images reconstructed by each algorithm enables verifying that they are of statistically indistinguishable quality. We therefore make sure that the chosen stopping rule guarantees that the achieved image quality is independent of the minimization algorithm. Such tests can obviously be carried out for several values of the stopping rule (for example, low, medium and high reconstruction qualities) to evaluate the performance of different algorithms in various situations.

The methodology presented above was used to compare four minimization algorithms. Here, we only report the key elements of this study, and the reader is referred to [HAM 11] for further details on the methodology and the results. The four algorithms were selected because of their good anticipated performance with a large-size nonlinear problem such as reconstruction in XRCT. Two of them were "generic": the first, L-BFGS-B [ZHU 97], is a limited-memory quasi-Newton method modified so as to integrate bound constraints; the second, IPOPT [WAC 06], is an interior point method based on sequential quadratic programming that can use a quasi-Newton type of approximation of the Hessian or of the Lagrangian. The other two methods were specifically developed for tomography: the first, SPS [ERD 99], operates on quadratic functions surrogate to convex criterion J; the second, TRIOT [AHN 06], is based on the same principle of quadratic surrogate functions, but uses an ordered subsets-based approach modified to preserve the convergence of the procedure. Comparisons were performed in a 2D framework with the monoenergetic model, for simplicity reasons and to limit the computation time. We advance the hypothesis that results obtained in this framework remain valid in more complex situations (3D geometry, polyenergetic model). Our results have shown that, when the quality of the desired reconstruction is low to moderate, specific algorithms, and notably TRIOT, have a slight advantage. When high quality reconstructions are desired, generic algorithms, and particularly L-BFGS-B, have a significant advantage. This behavior is illustrated in section 1.4.1. The practical objective of the works being to visualize small structures with accuracy, and therefore with high image quality, we selected the L-BFGS-B algorithm.

# 1.3.3.2. Minimization procedure

With 3D clinical data, directly using the L-BFGS-B algorithm to minimize criterion J is not an option, both due to the huge size of array A and to the corresponding computation time. However, for our vascular imaging application, the objective is to accurately visualize very small structures of interest. This leads naturally to decompose the object into a region of interest, that contains the structures to reconstruct with accuracy, and a background containing the rest of the object that does not need to be accurately imaged. From an algebraic point of view, such a decomposition is readily obtained by partitioning  $\mu$  into  $\mu_{\rm RI}$  (region of interest) and  $\mu_{\rm AP}$  (background), and by partitioning A correspondingly. The extension of this decomposition to the cases of thick rays and polyenergetic model does not present any difficulty.

At the methodological level, it is possible, under certain technical conditions, to estimate  $\mu_{\text{RI}}$  when  $\mu_{\text{AP}}$  is unknown [COU 08]. However, practical experience shows that the reconstruction of  $\mu_{\text{RI}}$  is both faster and of better quality when a reasonable estimation  $\mu_{\text{AP}}$  is available. Therefore, we consider that quantities  $\mu_{\text{RI}}$  and  $\mu_{\text{AP}}$  must be both reconstructed. The central issue is the impact on the estimation of  $\mu_{\text{RI}}$  of errors in  $\mu_{\text{AP}}$  such as a coarse resolution and/or a high estimation error. To accurately estimate  $\mu_{\text{RI}}$  while dedicating minimum effort in the determination of  $\mu_{\text{AP}}$ , several procedures can be envisioned; among these, we may single out (1) the prior

determination of  $\mu_{AP}$  either from images provided by the tomograph or by iterative reconstruction at a coarse resolution or with low quality; (2) the joint or sequential estimation of  $\mu_{RI}$  and  $\mu_{AP}$ .

These issues are discussed in detail in [HAM 10], and it appears that the quality of  $\mu_{AP}$  has little impact on the quality of the reconstruction of  $\mu_{RI}$ . This point is illustrated in section 1.4.2. Therefore, the preferred option is to determine  $\mu_{AP}$  once and for all, prior to the reconstruction of  $\mu_{RI}$ . The easiest way to obtain this prior estimation is to use reconstructions produced by the tomograph, if these are available. By default, reconstruction by analytical methods, or by iterative methods with a coarse resolution (typically one quarter of that used for  $\mu_{RI}$ ) and a low quality, can be used. In all cases, a significant reduction in the volume of computation and in memory requirements is obtained : the prior estimate of  $\mu_{AP}$  is either available, or computed analytically, or rebuilt iteratively in a small number of iterations, the projection matrix being of small size. The projection matrix used to reconstruct  $\mu_{RI}$  is also of small size, and this set of features makes it possible to implement the procedure on nonspecialized, if not modest, computer equipment.

#### **1.3.4.** Implementation of the reconstruction procedure

Finally, the method we have developed consists of minimizing the criterion  $J = J_{\rm V} + \lambda J_{\rm P}$  where  $J_{\rm V}$  and  $J_{\rm P}$  are respectively defined by [1.9]–[1.11] and [1.12]-[1.13]. The complexity of the data collection geometry (multiple detectors arrays, helical acquisition mode, flying focal spot, etc.) makes the construction of matrix A difficult and tedious, without being all the more overwhelming. For questions of computation speed, we have opted for prior construction over one rotation and storage of the projection operator. The formulation we have developed brings forward the possibility of systematically accounting for the thickness of rays in order to avoid artifacts produced in high-resolution reconstructions by the hypothesis of point detectors, and if necessary to use polyenergetic modeling of attenuation phenomena. With regard to the precise shape of penalty term  $J_{\rm P}$ , we have selected a second-order neighborhood within each slice, and a first-order neighborhood across slices (ten-neighbor model). This choice is justified by the anisotropy of the representation of  $\mu$ , the thickness of the slices being generally much larger than the pixel size in each slice. Finally, the decomposition of the object into background and region of interest is systematically used, because it corresponds to the characteristics of the reconstruction problem while being necessary to implement the method with the available computational resources. In most cases,  $\mu_{\rm AP}$  is determined from the images provided by the tomograph, and the expression of criterion J must be modified to involve  $\mu_{\rm RI}$  only. To perform the minimization, the L-BFGS-B algorithm is used, in accordance with the results of the comparison of the optimization algorithms and the targeted reconstruction quality. In addition, the use of this algorithm ensures in a simple and rigorous manner the positivity of the estimate of  $\mu_{RI}$ .

As for most inversion methods, the proposed reconstruction procedure requires the value of several parameters to be set, notably the number of rays per detector, the discretization stepsize of the spectrum of the X-ray source for the polyenergetic model, the value of the stopping rule, the value of hyperparameters  $\lambda$  and  $\delta$ . The unsupervised determination of such parameters is an important and difficult problem in inversion; conceptually, such an approach could be envisaged here for a number of quantities (notably  $\lambda$ ). However, from a practical point of view, this approach is to be ruled out due to the size of the reconstruction problem and to the amount of computation it would require. In addition, empirical determination of these parameters, according to the targeted quality objectives, does not raise major problems because of the availability of numerical and physical phantoms, and of the possibility of assessing the quality of reconstructions by experts. Thus, we have determined that five to ten rays per detector are largely sufficient to avoid artifacts during the reconstructions of  $\mu_{\rm RI}$ , and that a decomposition of the spectrum of the X-ray source into pulse components and smooth-variation curve make it possible to limit to ten the number of elements in the sums appearing in [1.11]. The other quantities may be set up by a calibration-based procedure to be carried out for each acquisition protocol.

## 1.4. Results

The objective of this section is not to provide experimental justification of the choices on which the method we have developed is based, but, more modestly, to support and illustrate by typical results some of the assertions made in the preceding paragraphs.

#### **1.4.1.** Comparison of minimization algorithms

In order to compare the four selected minimization algorithms, we have implemented the procedure described in section 1.3.3.1. It requires, as a first step, the definition of quality metrics of the reconstructed images and the design of random phantoms enabling the calculation of these metrics. Among the desired characteristics in the reconstructed images, we have retained the resolution, and the ability to distinguish low contrasts. It should be mentioned that a third commonly used characteristic, the residual noise affecting the reconstructed image, is strongly linked to the previous characteristics and therefore has not been retained. To evaluate these characteristics, we have used a 2D numerical phantom (220 mm sidelength, discretization on a  $512 \times 512$  grid) composed of a homogeneous medium with the

the attenuation of soft tissues, in which eight contrast gauges (disks of 10 mm in diameter whose attenuation value is close to that of the background) and eight resolution gauges (strongly attenuating disks of diameters ranging between 0.5 and 3 mm) are inserted. The random character of the phantom is achieved by partitioning it into 24 fixed cells, and by randomly placing 16 gauges in these 24 cells. An instance of such a phantom is represented in Figure 1.1. The quality metrics of the reconstructed images are defined as follows: for the ability to distinguish low contrasts, detection of contrast gauges is performed by matched filtering in the eight cells containing a contrast gauge and in the eight empty cells; the metric is equal to the proportion of accurate results. For resolution, the "modulation transfer function" (MTF) is used [DRO 82, HAM 10], which can be interpreted as a 1D frequency representation of the 2D impulse response of the reconstruction procedure, under the assumptions of spatial invariance and rotational symmetry. The impulse response is obtained by simple least-squares estimation from the original phantom and the reconstructed image, and the resolution metric is defined as twice the value of the normalized frequency for which the amplitude of the MTF has decreased by 40% relative to its maximum value.



Figure 1.1. Instance of the 2D random numerical phantom used for the comparison of optimization algorithms (based on [HAM 11]). Axis scales in mm

The data were obtained by simulating a current generation tomograph, in 2D axial mode. The geometrical parameters of a Siemens SOMATOM 16 scanner were used, the projection process being simulated by the SNARK09 software [DAV 09] taking into account the polyenergetic character of the source. To avoid the inverse crime [KAI 07] as much as possible, an analytical parameterization of the phantoms was used, and, in order to better model the propagation of X-rays in the medium, the

detectors were separated into several cells whose photon counts were then combined. To carry out the comparison, 50 independent instances of the random phantom were drawn, and the random phenomena involved in the data formation process were also independent of each other.

The criterion used for the reconstructions was based on a monoenergetic data formation model. The difference between the model used to generate data and that used to perform the reconstructions corresponds to the most frequent practical situation. In a preliminary step, the hyperparameter values were calibrated from a data set different from those used for the comparisons, and three values of the stopping rule, corresponding visually to a moderate, good and very good quality of the reconstructed image, were determined. The criterion J was then minimized with the four methods mentioned in section 1.3.3.1, for each of the 50 datasets. In order to make the comparison as fair as possible, all methods were initialized in an identical fashion (image obtained after 20 iterations of the non-convergent OS-SPS method, itself initialized by an analytical reconstruction), and the elements having a significant impact on the numerical performance of algorithms were implemented in the same way. The used performance measure was the number of projections and backprojections required to satisfy the stopping rule, this measure being directly related to the total volume of computation required to perform the reconstruction. For each of the three values of the stopping rule, a comparison of the quality measures of the reconstructed image was made, to ensure lack of any statistically significant difference from one algorithm to another.

The results are synthesized in the diagram of Figure 1.2, which presents the performance measure and its standard deviation (in positive values for better readability) for each algorithm as a function of the value of the stopping rule. These results show that, for a high value of the stopping rule (low reconstruction quality), the TRIOT method is the most effective, closely followed by L-BFGS-B and SPS, the IPOPT algorithm being much slower. However, for lower values of the stopping rule (better image quality), algorithms specifically developed for tomography are outperformed by generic methods IPOPT and L-BFGS-B, the latter presenting the best overall performance. This behavior can be explained by the fact that SPS and TRIOT are essentially gradient algorithms, and that their asymptotic convergence is therefore slower than that of quasi-Newton methods such as IPOPT and L-BFGS-B. A low variance of performance measurements can be observed, except for IPOPT. This behavior could be attributed to the particular nature of this algorithm (interior point method). In addition, it should be mentioned that for each of the three selected stopping rule values, the quality of the reconstructions provided by the four algorithms showed no statistically significant difference. These results are in agreement with the behavior defined in section 1.3.3.2 and justify our choice of the L-BFGS-B method to carry out the minimization of criterion J in our reconstruction methods.

#### 1.4.2. Using a region of interest in reconstruction

Our purpose is to illustrate the performance of the region of interest-based minimization procedure described in section 1.3.3.2. Specifically, we are interested in the total reconstruction time as well as in the quality of the reconstructed region of interest for various acquisition modes of the background. In the previous paragraph, we presented a way to evaluate and to compare the quality of images reconstructed by several methods. Here, this technique cannot be directly applied, because the images to compare do not result from the same estimator, since the background varies from one reconstruction to another. This is why the quality of regions of interest is evaluated in a more limited way, on the one hand by visual inspection for the same setting of hyperparameters, and, on the other hand, by calculating the resolution of regions of interest reconstructed for a same estimation variance, these being obtained by an adequate adjustment of the hyperparameters.



**Figure 1.2.** Comparison of the average performance of the four retained algorithms, depending on the value of the stopping rule (from [HAM 11]). Algorithms specifically developed for tomography are more effective for high stopping criteria (low quality of reconstruction), but are outperformed by generic methods L-BFGS-B and TRIOT for stricter stopping rules

The tests were carried out on data simulated in a 2D framework. The phantom (sidelength of 200 mm, discretization on a  $1240 \times 1240$  grid) being composed of a homogeneous medium with the attenuation of soft tissues, in which several objects were placed: first, in the peripheral zone (background), 10 strongly attenuating disks were included; their function was to allow us to assess the impact of an inaccurate representation of such objects on the reconstruction quality of the region of interest. Two versions of the phantom, in which the attenuation of the disks was either that of iron or that of aluminum, were used. Second, the central part (region of interest, with a 30 mm sidelength and  $186 \times 186$  grid size) comprised two types of objects,

respectively moderately and strongly attenuating, arranged in such manner that resolution and estimation variance could be simply assessed. The phantom is represented in Figure 1.3, and the corresponding data were generated using the same procedure as that described in section 1.4.1.



b) Central part, corresponding to the region of interest, of size  $186 \times 186$ 

**Figure 1.3.** Phantom used to test the approach to reconstruction of the medium by decomposition in background and region of interest (based on [HAM 10]). The gray levels correspond to an attenuation in Hounsfield units, on a [-1000,3000] scale

From these data, the background was reconstructed first, with an analytical method (filtered backprojections) at full resolution, and second, by iteratively minimizing the criterion *J* under a monoenergetic assumption, using the L-BFGS-B algorithm, with resolutions of 1,240, 480, 240, 160, 120, 80 and 40. The quality of reconstruction not being critical, the stopping rule was set to a high value. For each resolution of the background, the region of interest was then reconstructed, still with the proposed iterative method under the monoenergetic assumption, under the following conditions: for visual assessment, the regularization parameter was set to a low value, so as to bring forward the artifacts related to the imperfect representation of the background, and the value of the stopping rule was set to an intermediate value (good quality of reconstruction). For the evaluation of the resolution, the regularization parameter was adjusted for each reconstruction. There again, the value of the stopping rule was set to an intermediate value common to all reconstructions. As in the previous paragraph, the resolution was evaluated from the MTF of the reconstruction [DRO 82, HAM 10].



**Figure 1.4.** Analysis of regions of interest obtained with a low common value of the regularization parameter (based on [HAM 10]). Each thumbnail represents the difference between the exact region of interest and the region of interest reconstructed with a particular background, on a [-250, 250] scale in Hounsfield units. The background disks presented the attenuation characteristics of iron. There is an almost complete absence of artifacts for background resolutions higher than 120

Figure 1.4 presents the results of the reconstruction of the region of interest for a low value of the regularization parameter when background discs have the

attenuation characteristics of iron. In order to better appreciate the presence of artifacts, each thumbnail represents the difference between the exact region of interest and the reconstructed region of interest with a particular background. When the background is reconstructed iteratively, lack of artifacts can be observed for resolutions higher than 120, which corresponds to a factor greater than 10 between the sidelengths of pixels of the background and of the region of interest. Significant artifacts can also be observed when the analytically reconstructed background is used at full resolution. When the background disks have the attenuation characteristics of aluminum (less attenuating than iron), the results of the region of interest reconstructed background where no artifacts are observed [HAM 10].

MTFs corresponding to the reconstructions of the region of interest with backgrounds obtained iteratively with different resolutions are presented in Figure 1.5. The background disks had the attenuation characteristics of iron. It can be observed that there is no significant loss of resolution in the region of interest for background resolutions greater than 120. Similar results were obtained when background disks had the attenuation characteristics of aluminum [HAM 10]. These results are thus consistent with the visual inspection of reconstruction artifacts. Finally, our experiments showed that the total volume of computation required to reconstruct the region of interest is lower when the background is obtained analytically, then grows according to the resolution chosen to iteratively reconstruct the background.

These results indicate that the quality of the reconstruction of the region of interest is relatively insensitive to that of the background. In most cases, an analytical reconstruction of the background with full resolution of the background is sufficient, except in the presence of very strongly attenuating objects (e.g. prostheses with stainless steel components). In this case, iterative reconstruction with low resolution (factor up to 10 on the pixel sidelength) should be used, which results in a substantial increase in the total amount of computation. These simulation results have been confirmed by tests on real data carried out with a physical resolution phantom (CTP 528 section of the Catphan600© phantom, see [HAM 10] for full results). This validates, at least in a 2D context, the reconstruction approach by decomposition of the object into background and region of interest, and confirms that it is unnecessary to carry out the joint, simultaneous or sequential estimation of  $\mu_{AP}$  and  $\mu_{RI}$ .

### **1.4.3.** Consideration of the polyenergetic character of the X-ray source

The results reported in Figure 1.5 support the choices made in section 1.3.3 regarding the algorithm and the minimization procedure, at least in the framework adopted to perform those tests (2D geometry, monoenergetic model). We hypothesize that these results remain valid in more complex situations (3D geometry,

polyenergetic model). The subject of this section is to briefly illustrate the effect of accounting for the polyenergetic nature of the source, first with simulated data in 2D, then with real data in 3D.



**Figure 1.5.** *MTF* of regions of interest reconstructed with backgrounds obtained iteratively at different resolutions (according to [HAM 10]). Background disks with the attenuation characteristics of iron. It can be observed that there is no significant loss of resolution for background resolutions greater than 120. Similar results were obtained when the background disks had the attenuation characteristics of aluminium

#### 1.4.3.1. Simulated data in 2D

Data were generated using the same phantom (disks located in the background: attenuation characteristics of iron) and the same simulation process as for the evaluation of the approach by region of interest. This choice is justified by the presence of strongly attenuating objects within the phantom, and by the fact that the polyenergetic character of the X-ray source is accounted for in the projection simulation process. First, the background was reconstructed analytically, on a 512 × 512 grid. Second, the region of interest was estimated by the proposed method, under the monoenergetic and polyenergetic hypotheses, with the same resolution as that of the background (size  $80 \times 80$ ). In both cases, the hyperparameters were empirically adjusted, and the stopping rule was set to a low value (very good quality of reconstructed image) such as to bring forward the differences between methods.



**Figure 1.6.** Results of the reconstruction of the region of interest under monoenergetic and polyenergetic hypotheses. The gray levels correspond to the attenuation at 70 keV expressed in Hounsfield units, on a [-1000,3000] scale. Under a monoenergetic hypothesis, it can be observed that there are many artifacts, and several structures present in the region of interest are not reconstructed. Under a polyenergetic hypothesis, these defects have disappeared almost completely

An example of reconstructions of regions of interest is presented in Figure 1.6. It can be observed that under a monoenergetic hypothesis, the reconstructed image presents visible defects: shadow areas due to the disks of the background, significant

underestimation of the attenuation at the center of the image, absence of certain structures. These defects have almost completely disappeared from the reconstructed object when accounting for the polyenergetic character of the source during the reconstruction of the region of interest. This improvement is achieved at the cost of an increase in the volume of computation by a factor ranging between 2 and 3.

These results, and the many other tests we have carried out, indicate that accounting for the polyenergetic character of the X-ray source in the reconstruction method leads to a significant improvement in the quality of reconstructions, in the presence of strongly attenuating objects, both in the background and in the region of interest. It should be underlined that this improvement depends on the accuracy of the polyenergetic attenuation model, which in practice, can be difficult to assess, or even questionable.

#### 1.4.3.2. Real data in 3D

The results presented here were obtained from the multimodality vascular phantom schematically represented in Figure 1.7. It consists of a tube, simulating a blood vessel, with two calibrated stenoses, one of which is located inside a *stent*. This setup is included in a block of agar simulating soft tissue, and a contrast agent can be injected into the tube. The phantom was imaged with a clinical tomograph (Siemens SOMATOM 16) using the protocol used for peripheral vascular examinations (16 arrays of 672 detectors each, helical mode, 1,160 projections per rotation, flying focal spot). The reconstructions were performed by the tomograph in accordance with this protocol ( $512 \times 512$  images, 0.7 mm spaced slices) on a field of view of 15 cm in diameter.



Figure 1.7. Multimodality vascular phantoms used to generate the data. The stent is placed around one of the two stenoses

Reconstructions were carried out with the proposed method, under monoenergetic and polyenergetic hypotheses, using the raw data collected by the tomograph. It should be emphasized that these "raw" data have undergone extensive preprocessing, part of which aiming to limit the artifacts related to the polyenergetic character of the X-ray source. Therefore, this casts a doubt on the accuracy of the polyenergetic attenuation model that we use. The chosen region of interest had a sidelength of 3.25 cm and was discretized on a  $128 \times 128 \times 94$  grid to obtain the same voxel size as that of the reconstruction generated by the tomograph. The background was reconstructed analytically by the tomograph, on a full field of view 50 cm in diameter ( $512 \times 512$  grid). The hyperparameters were empirically adjusted so as to achieve the same estimation variance as that of images provided by the tomograph, and the stopping rule was set to an average value to obtain a satisfactory image quality without unduly increasing the volume of computation.



**Figure 1.8.** Results of the reconstruction of the region of interest. Gray levels correspond to the attenuation at 70 keV expressed in Hounsfield units, on a [-200,1100] scale. The proposed reconstructions present a significant improvement

An example of reconstruction is presented in Figure 1.8. The selected slice corresponds to the extremity of a *stent* fitted with distal markers. It can be observed that the method we have developed produces a significant improvement over the image provided by the tomograph, notably with regard to the presence of shadow

areas and to the size of the strongly attenuating markers. However, reconstructions under monoenergetic and polyenergetic hypotheses appear almost identical.

A more complete evaluation of these methods was made from the reconstructions of eight vascular phantoms. After result anonymization and random ordering, several clinically significant parameters were evaluated in two ways: by semi-automatic measures resulting from a clinical image analysis software, and by manual measurements independently carried out by two physicians. Statistical analysis confirmed that iterative reconstructions provide more accurate indications about the shape and size of the *stent*. In addition, this analysis showed that, under a polyenergetic hypothesis, intra-*stent* stenosis measurements are more accurate (in the statistical sense) than in the other two methods. These results thus confirm the conclusions drawn from the 2D simulation results.

# 1.5. Conclusion

Our goal was to develop a 3D reconstruction method in XRCT capable of processing clinical data for vascular imaging. The difficulties related to the modeling of the data formation process and to the size of the corresponding estimation problem have been overcome, at least partially, by developing a parsimonious representation of the projection operator, by decomposing the medium to reconstruct into background and region of interest, by developing an optimization procedure tailored to this decomposition and by carefully selecting the minimization algorithm. The last difficulty mentioned in the introduction, that is to say, problems to access the raw data, has been overcome by combining a certain degree of collaboration with the manufacturer and a significant dose of reverse engineering in order to determine the precise structure of the datafiles. The method we have developed provides room for many improvements aiming to accelerate computations particularly with regard to the representation and the coding of the projection / backprojection operator, and to the implementation of the optimization procedure (massive parallelization for example). Nevertheless, the results indicate that, for some specific applications, methods such as those presented here are of real practical interest, because they provide reconstructed images of a significantly higher quality than the images produced by clinical apparatus.

More generally, image reconstruction is a typical example of an ill-posed inverse problem, and it is interesting to analyze the nature of the difficulties and the manner to overcome them with regard to the elements presented in Demoment's synthesis article [DEM 89], which reflects the state of our knowledge a little more than 20 years ago. Demoment's article focuses on the methodological aspects of solving image reconstruction and restoration problems, essentially under the assumption of linearity of the data formation model. The article insists on the notion of ill-posed inverse problem, on various regularization techniques and on the question of the determination of the value of hyperparameters. The article also deals, in a less detailed manner, with the case of nonlinear data formation models such as those encountered in diffraction tomography, and with the difficulties of the practical computation of the solution.

With regard to 3D reconstruction in XRCT, the method presented here, as well as recent literature, clearly indicate that the ill-posed character of the problem, and the regularization techniques necessary for its solution, are now recognized and taken into account by the research community. The question of the specification of the value of hyperparameters does not pose any major practical difficulty, due to the existence of specific acquisition protocols which allow us to calibrate reconstruction methods. The interest of the community seems to have shifted toward the application of these methods to real large-size datasets, along with the problems of computation time and implementation arising therefrom. The development of methods usable in practice involves the harmonious choice of a data formation model, of a minimization procedure for the resulting criterion, and of specific optimization algorithms. The method that we have presented is a – quite imperfect – attempt pointing in this direction. A significant proportion of recent works in 3D XRCT also follows this path [KAC 11].

This situation does not seem limited to XRCT. The same trend can be perceived in emission tomography and diffraction tomography, but with certain nuances such as the attachment of certain communities to the automatic determination of the value of hyperparameters [BER 97]. This can be interpreted as the result of a widespread recognition of the methodological issues synthesized in [DEM 89]. This recognition has made it possible to address problems hampered by the presence of nonlinearities in the data formation process or by their ill-conditioning, with the frequent additional difficulty of a large volume of data. This explains the important place taken by implementation issues. This brief analysis thus brings forward the impact that methodological studies may have, some 20 years later, on the approach to various problems and the development of practical solutions.

# 1.6. Acknowledgments

The works presented here began in 2006 and have benefited from the contribution of many collaborators. In particular, we should mention Jean-Pierre Dussault (University of Sherbrooke, Canada), Gilles Soulez, Guy Cloutier, Gilles Beaudoin and Sofiane Hadjadj (*Centre de recherche du Centre hospitalier de l'Université de Montréal* (CRCHUM) – Research Center of the University of Montreal Health Centre, Canada). Students and postdoctoral fellows involved in the project have had a decisive impact on its evolution and on the results obtained. We should emphasize the contributions of Malik Dussaud, David Gendron, Benoît Hamelin and Saul Pérez.

Research partially funded by the Natural Sciences and Engineering Research Council of Canada (Discovery grant no. 138417-06 and Collaborative health research project no. 323572-06), and by the Canadian Institutes of Health Research (Collaborative health research project no. 323572-06).

## 1.7. Bibliography

- [AHN 06] AHN S., FESSLER J.A., BLATT D., et al., "Convergent incremental optimization transfer algorithms: Application to tomography", *IEEE Transactions on Medical Imaging*, vol. 25, no. 3, pp. 283–296, March 2006.
- [ALV 76] ALVAREZ R.E., MACOVSKI A., "Energy-selective reconstructions in X-ray computerized tomography", *Physics in Medicine and Biology*, vol. 21, no. 5, pp. 733–744, September 1976.
- [AMB 73] AMBROSE J., HOUNSFIELD G.N., "Computerized transverse axial tomography", *British Journal of Radiology*, vol. 46, no. 542, pp. 148–149, February 1973.
- [BER 97] VAN DEN BERG P.M., KLEINMAN R.E., "A contrast source inversion method", *Inverse Problems*, vol. 13, no. 6, pp. 1607–1620, December 1997.
- [BOA 04] BOAS D.A., DALE A.M., FRANCESCHINI M.A., "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy", *NeuroImage*, vol. 23, pp. S275–S288, 2004.
- [BOU 93] BOUMAN C.A., SAUER K.D., "A generalized Gaussian image model for edge-preserving MAP estimation", *IEEE Transactions on Image Processing*, vol. 2, no. 3, pp. 296–310, July 1993.
- [COU 08] COURDURIER M., NOO F., DEFRISE M., *et al.*, "Solving the interior problem of computed tomography using *a priori* knowledge", *Inverse Problems*, vol. 24, no. 6, December 2008.
- [DAV 09] DAVIDI R., HERMAN G.T., KLUKOWSKA J., "SNARK09: A programming system for the reconstruction of 2D images from 1D projections", 2009. Available at: www.snark09.com.
- [DEL 98] DELANEY A.H., BRESLER Y., "Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography", *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 204–221, February 1998.
- [DEM 89] DEMOMENT G., "Image reconstruction and restoration: Overview of common estimation structures and problems", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-37, no. 12, pp. 2024–2036, December 1989.
- [DEM 01] DE MAN B., NUYTS J., DUPONT P., et al., "An iterative maximum-likelihood polychromatic algorithm for CT", *IEEE Transactions on Medical Imaging*, vol. 20, no. 10, pp. 999–1008, October 2001.
- [DEM 04] DE MAN B., BASU S., "Distance-driven projection and backprojection in three dimension", *Physics in Medicine and Biology*, vol. 49, no. 11, pp. 2463–2475, June 2004.

- [DRO 82] DROEGE R.T., MORIN R.L., "A practical method to measure the MTF of CT scanners", *Medical Physics*, vol. 9, no. 5, pp. 758–760, September 1982.
- [ERD 99] ERDOĞAN H., FESSLER J.A., "Ordered subsets algorithms for transmission tomography", *Physics in Medicine and Biology*, vol. 44, no. 11, pp. 2835–2851, November 1999.
- [FES 94] FESSLER J.A., "Penalized weighted least-squares image reconstruction for positron emission tomography", *IEEE Transactions on Medical Imaging*, vol. 13, no. 2, pp. 290– 300, 1994.
- [GEN 08] GENDRON D., GOUSSARD Y., BEAUDOIN G., et al., "Reconstruction of real tomographic data using algebraic methods", *IEEE International Conference of the Engineering in Medicine and Biology Society*, Vancouver, BC, Canada, pp. 2717–2720, August 2008.
- [GEN 11] GENDRON D., GOUSSARD Y., HAMELIN B., et al., "Reconstruction of real tomographic data using algebraic methods", *Proceedings of the 11th International Meeting* on Fully 3D Image Reconstruction, Potsdam, Germany, pp. 343–346, July 2011.
- [HAM 10] HAMELIN B., GOUSSARD Y., DUSSAULT J.-P., et al., "Design of iterative ROI transmission tomography reconstruction procedures and image quality analysis", *Medical Physics*, vol. 37, no. 9, pp. 4577–4589, September 2010.
- [HAM 11] HAMELIN B., GOUSSARD Y., DUSSAULT J.-P., et al., "Performance comparison on nonlinear solvers for statistical reconstruction in transmission tomography", *IEEE Transactions on Medical Imaging*, October 2011.
- [HER 71] HERMAN G.T., ROWLAND S.W., "Resolution in ART: An experimental investigation of the resolving power of an algebraic picture reconstruction technique", *Journal of Theoretical Biology*, vol. 33, pp. 213–223, 1971.
- [HER 73] HERMAN G.T., ROWLAND S.W., "Three methods for reconstructing objects from X-rays: a comparative study", *Computer Graphics and Image Processing*, vol. 2, pp. 151–178, 1973.
- [HER 76a] HERMAN G.T., LENT A., "A computer implementation of a Bayesian analysis of image reconstruction", *Information and Control*, vol. 31, no. 4, pp. 364–384, August 1976.
- [HER 76b] HERMAN G.T., LENT A., "Quadratic optimization for image reconstruction. I", Computer Graphics and Image Processing, vol. 5, pp. 319–332, 1976.
- [HER 79] HERMAN G.T., HURWITZ H., LENT A., et al., "On the Bayesian approach to image reconstruction", *Information and Control*, vol. 42, pp. 60–71, 1979.
- [HOU 73] HOUNSFIELD G.N., "Computerized transverse axial scanning (tomography). 1. description of system", *British Journal of Radiology*, vol. 46, no. 552, pp. 1016–1022, December 1973.
- [KAC 11] KACHELRIESS M., RAFECAS M., (eds.), Proceedings of the 11th International Meeting on Fully 3D Image Reconstruction, Potsdam, Germany, July 2011.

- [KAI 07] KAIPIO J., SOMERSALO E., "Statistical inverse problems: discretization, model reduction and inverse crimes", *Journal of Computational and Applied Mathematics*, vol. 198, no. 2, pp. 493–504, 2007.
- [LEV 87] LEVITAN E., HERMAN G.T., "A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography", *IEEE Transactions on Medical Imaging*, vol. MI-6, pp. 185–192, September 1987.
- [SAU 93] SAUER K.D., BOUMAN C.A., "A local update strategy for iterative reconstruction from projections", *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 534–548, February 1993.
- [TRI 10] TRILLON A., Amélioration de méthodes de reconstruction de défauts à partir de données issues de capteurs à courants de Foucault, thesis, École centrale de Nantes, France/École polytechnique de Montréal, Québec, Canada, October 2010.
- [VAU 11] VAUTRIN D., VOORONS M., IDIER J., et al., "Seismic imaging of transmission overhead line structure foundations", Computational Imaging IX (IS&T/SPIE Symposium on Electronic Imaging), San Francisco Airport, CA, USA, January 2011.
- [WAC 06] WÄCHTER A., BIEGLER L.T., "On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming", *Mathematical Programming*, vol. 106, no. 1, pp. 25–57, May 2006.
- [ZHU 94] ZHUANG W., GOPAL S.S., HEBERT T.J., "Numerical evaluation of methods for computing tomographic projections", *IEEE Transactions on Nuclear Sciences*, vol. 41, no. 4, pp. 1660–1665, August 1994.
- [ZHU 97] ZHU C., BYRD R.H., LU P., et al., "Algorithm 778. L-BFGS-B: Fortran subroutines for Large-Scale bound constrained optimization", ACM Transactions on Mathematical Software, vol. 23, no. 4, pp. 550–560, 1997.