Metal Artefact Reduction by Least-Squares Penalized-Likelihood Reconstruction with a Fast Polychromatic Projection Model

B Hamelin,[1] Y Goussard,[1] J P Dussault,[2] G Cloutier,[3] G Beaudoin[3] and G Soulez[3]

[1] Institut de génie biomédical, École Polytechnique de Montréal, Montréal, Québec, Canada.
[2] Département d'informatique, Université de Sherbrooke, Sherbrooke, Québec, Canada.
[3] Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, Québec, Canada.

E-mail: benoit-2.hamelin@polymtl.ca

Abstract. We consider penalized-likelihood reconstruction for X-ray computed tomography of objects that contain small metal structures. To reduce the beam hardening artefacts induced by these structures, we derive the reconstruction algorithm from a projection model that takes into account the photon emission spectrum and nonlinear variation of attenuation to photon energy. This algorithm requires excessively long runtimes to attain convergence. We contribute refinements in the representation of the object and the modelling of the photon emission spectrum aimed at reducing the runtime.

These refinements yield an algorithm for which the reconstruction runtime is reduced by at least one order of magnitude on each considered dataset. For real data that were preprocessed for metal artefact reduction, the enhanced polychromatic algorithm provides the best observed image quality for objects with large metal structures. For objects with small metal structures, the gains are less significant with respect to a penalized-likelihood algorithm derived from a monochromatic projection model.

PACS numbers: 87.57cp,87.57nf,87.57nt,87.57Q-

Submitted to: *Phys. Med. Biol.*

1. Introduction

We take interest in metal artefact reduction in computed tomography (CT) reconstruction. The application subcontext is that of vascular imaging in the presence of a metal stent.

For the usual reconstruction algorithms, metal objects are responsible for severe beam hardening artefacts. This phenomenon is caused by the nonlinear variation of the attenuation of X photons with respect to photon energy, coupled to the nonuniformity of the polychromatic X-ray beam used for scanning the body. Usual reconstruction algorithms based on a monochromatic projection model, such as filtered backprojection, do not account for either. Therefore, predicted photon counts tend to be inconsistent with measured photon counts. This induces streak artefacts around metal objects.

Another important source of artefacts is measurement noise, which may be attributed to two main causes. First, X-ray propagation through the object is best described in terms of particle interaction phenomena, which confers a stochastic aspect to the photon counts. Second, tomographic data suffers from many alterations that induce systematic artefacts in reconstructed images. These alterations stem from limitations of the hardware and projection model and include off-focal radiation and detector effects such as afterglow and crosstalk (LaRivière et al., 2006). Most modern scanners implement various data preprocessing techniques to reduce these artefacts, many of which can amplify data noise (LaRivière and Vargas, 2006).

Recent literature is rife with contributions concerned with metal artefact reduction. Many of these contributions realize the reduction by preprocessing the sinogram, to which a common analytical reconstruction algorithm (e.g. filtered backprojection, Feldkamp-Davis-Kress reconstruction, etc.) is applied. The preprocessing is either a subtraction of the contribution of the metal structures (which are reconstructed apart from the rest of the object) completed by interpolation (Rinkel et al., 2008; Oehler et al., 2008), or a nonlinear mapping to a monochromatic sinogram representation (Mou et al., 2008; LaRivière et al., 2006). However, these algorithms are sensitive to measurement noise, as they do not account or compensate for it. Moreover, metal artefact reduction by sinogram subtraction and simplistic interpolation may induce other artefacts in reconstructed images (Müller and Buzug, 2009).

Statistical reconstruction algorithms (Fessler, 2000) are more robust to the data inconsistencies that cause these artefacts. They intrinsically take the stochastic aspect of the data formation process into account and can compensate for it through the introduction of prior information on the object. Oehler and Buzug (2007) used such a reconstruction procedure on a sinogram that had been preprocessed by a metal artefact method such as mentioned above.

In addition, the statistical framework for CT reconstruction can be adapted to projection models that describe the beam hardening effect. The development of algorithms in this framework around a polychromatic projection model has been the onus of many contributions since 2000. De Man et al. (2001) offered an algorithm hinged on a discrete representation of the source spectrum and a polyenergetic object representation mapped from a single attenuation image at a reference energy. Using the same model, we proposed an alternative reconstruction algorithm that was faster and more robust to data preprocessing (Hamelin et al., 2008). An approach similar to that of De Man et al. was described by Chueh et al. (2008), where the explicit source spectrum is replaced with spectral data gathered from blank scans. Elbakri and

Fessler (2002, 2003) presented an algorithm for beam hardening reduction, though not explicitly for metal imaging. This approach relies on the decomposition of the object into two material maps, so that it may be represented by a single material proportion image. This object representation is also used by O'Sullivan and Benac (2007), who solve it with a different algorithm.

The main criticism leveled against statistical reconstruction concerns its high computational cost. These methods tend to require a lot of memory to run and a lot of runtime to converge. However, we rest on the hypothesis that the gain in image quality with respect to faster analytical algorithms can translate into clinical benefits. As such, the primary goal of this paper is to present strategies to reduce the runtime of polychromatic statistical reconstruction algorithms. While these strategies are specifically designed around the projection model of De Man et al. (2001), some may be ported to other similar setups. Namely, runtime reduction relies on two ideas. The first one is a targeted reconstruction approach, by which an initial reconstruction on a coarse image grid is used as the *background* to bootstrap the reconstruction of a small *region of interest* (ROI), represented on a finer grid. The second one is a concise discrete model of the source spectrum that takes the X photon emission phenomena into account. Along with the efficient algorithm of Hamelin et al. (2008) for solving the regularized polychromatic reconstruction problem, these enhancements to the model of De Man et al. yield a dramatic reduction of the computational cost without altering image quality.

The secondary goal is to compare the performance of some algorithmic variants over real datasets, both in terms of image quality and numerical efficiency. These datasets are obtained from the acquisition of a vascular phantom implanted with a metal stent, as well as some other metal objects to present various degrees of beam hardening at different points along the rotation axis. The data are preprocessed by the scanner software to reduce various families of artefacts, among which beam hardening artefacts. All algorithmic variants that we compare are based on projection models that do not take this undisclosed preprocessing into account. This comparison yields some insight in the performance of preprocessing at improving image quality with respect to polychromatic projection modelling.

The rest of the paper is organized as follows. Section 2 recalls polychromatic reconstruction by a penalized-likelihood approach. Section 3 discusses the refinements to the projection model that lead to a faster implementation. Section 4 presents the validation of the enhancements to the polychromatic projection model through the reconstruction of said simulation datasets. Section 5 compares reconstruction algorithms with metal artefact reduction over real datasets. Finally, Section 6 offers concluding remarks.

2. Polychromatic reconstruction

The following development is concerned with the polychromatic projection model of De Man et al. (2001), specially tailored for the reduction of metal artefacts (as opposed to that of sole beam hardening artefacts). Nonetheless, the acceleration techniques developed in Section 3 for this model and reconstruction problem may be ported to the models of Elbakri and Fessler (2003).

2.1. Projection model

The basic deterministic data formation model stems from the discretization of the Beer-Lambert law of X-ray attenuation (Kak and Slaney, 1987) that takes into account the nonuniform source emission spectrum:

$$\tilde{\boldsymbol{y}} = \int_0^{E_{\mathrm{kVp}}} b(E) \exp(-\mathbf{A}\boldsymbol{\mu}(E)) \mathrm{d}E \approx \sum_{k=1}^{K} b_k \exp(-\mathbf{A}\boldsymbol{\mu}^k). \tag{1}$$

Equation (1) introduces many symbols that will be used throughout: $\tilde{\boldsymbol{y}} \in \mathbb{R}^n$ represents the photonic intensity measured at each detector and projection angle; $\boldsymbol{b} = [b_1, b_2 \dots b_K]^{\mathrm{t}} \in \mathbb{R}^K$ is the source emission spectrum discretized over K equispaced energy values and scaled in order to properly approximate the integral over the continuous spectrum; $\boldsymbol{\mu}^k \in \mathbb{R}^m$ is the attenuation image of the object at energy value k; $\mathbf{A}$ is a $n \times m$ matrix that embodies the numerical integration of the image along each of the n raypaths. The difficulty with the derivation of a reconstruction algorithm from (1) is that K attenuation images must be computed. The corresponding system of nonlinear equations is woefully underdetermined. The core of the contributions in statistical methods for polychromatic reconstruction is a model for mapping a unique image to energy-dependant attenuation images.

The solution proposed by De Man et al. (2001) is to compute these K attenuation maps from a reference image $\boldsymbol{\mu}^0$ at a chosen energy E_0 through the decomposition of Alvarez-Macovski (Alvarez and Macovski, 1976). This approach expresses the attenuation coefficient of an object for any photon energy as the weighted sum of the contributions of the photo-electric effect and of Compton scattering to X-ray attenuation. These effects are assumed to depend only on the material, while the weights vary according to energy. Formally, $\boldsymbol{\mu}^k = \Phi_k \boldsymbol{\phi} + \Theta_k \boldsymbol{\theta}$, with $\Phi_k = (E_0/E_k)^3$ and $\Theta_k = f_{\mathrm{KN}}(E_k)/f_{\mathrm{KN}}(E_0)$, where E_0 is a reference energy level and $f_{\mathrm{KN}}(E)$ is the Klein-Nishina function (De Man et al., 2001).

The Alvarez-Macovski decomposition still represents the object with two attenuation maps. To reduce this representation to a single map, De Man et al. (2001) computed the $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ values by least squares for a small set of materials for which the energy to attenuation coefficient map is known. He used these values to set up empirical nonlinear maps $\phi(\boldsymbol{\mu}^0)$ and $\theta(\boldsymbol{\mu}^0)$ from attenuation at the reference energy to attenuation by photo-electric effect and by Compton scattering, respectively. The polychromatic projection model (1) may then be recast as

$$\tilde{\boldsymbol{y}} = \boldsymbol{p}(\boldsymbol{\mu}, \boldsymbol{b}) \stackrel{\mathrm{def}}{=} \sum_{k=1}^{K} b_k \exp\left\{-\mathbf{A}[\Phi_k \boldsymbol{\phi}(\boldsymbol{\mu}^k) + \Theta_k \boldsymbol{\theta}(\boldsymbol{\mu}^k)]\right\}, \tag{2}$$

where $\boldsymbol{\phi}(\boldsymbol{\mu}) : \mathbb{R}^m \to \mathbb{R}^m, \mu_i \mapsto \phi(\mu_i)$ and similarly for $\boldsymbol{\theta}(\boldsymbol{\mu})$.

2.2. Penalized-likelihood reconstruction

Hard structures in the image object lead to low photon counts. In this case, measurement uncertainty is dominated by photon propagation processes, which are best described by a Poisson distribution (LaRivière et al., 2006) with mean $\boldsymbol{p}(\boldsymbol{\mu}, \boldsymbol{b})$ (2). This statistical approach reduces reconstruction to the estimation of its parameter $\boldsymbol{\mu}$.

In this context, the maximum *a posteriori* (MAP) bayesian estimation framework yields a robust algorithm. This framework allows one to introduce known information

on the solution through the definition of a prior distribution. Considering Poisson-distributed photon counts and a prior distribution from the exponential family, the polychromatic reconstruction problem is expressed as

$$\min_{\boldsymbol{\mu}\geq 0} \sum_{i=1}^{n} [p_i(\boldsymbol{\mu}, \boldsymbol{b}) - \tilde{y}_i \ln p_i(\boldsymbol{\mu}, \boldsymbol{b})] + \lambda R(\boldsymbol{\mu}), \tag{3}$$

where $p_i(\boldsymbol{\mu}, \boldsymbol{b})$ denotes the i-th component of vector $\boldsymbol{p}(\boldsymbol{\mu}, \boldsymbol{b})$ and $R(\boldsymbol{\mu})$ is the penalty term, weighted by λ against the first term, which measure data adequation.

When dealing with real data, the preprocessing done by the scanner software contributes significantly to data noise. Data uncertainty is therefore not dominated anymore by photon propagation events. Similarly to Fessler (1994), who proposed this approach for precompensated data in positron emission tomography, Hamelin et al. (2008) used a Gaussian distribution of data uncertainty, which is a good approximation to the Poisson distribution for large photon counts. The reader is referred to Hamelin et al. (2008) for a full justification. With a Gaussian uncertainty model, the MAP estimator is the solution to the regularized nonlinear least squares problem

$$\min_{\boldsymbol{\mu}\geq 0} \frac{1}{2} \|\boldsymbol{y} + \ln \boldsymbol{p}(\boldsymbol{\mu}, \boldsymbol{\beta})\|^2 + \lambda R(\boldsymbol{\mu}). \tag{4}$$

The special case of $K = 1$ is of particular interest. Equation (1) then devolves in the monochromatic projection model, e.g. $\ln \boldsymbol{p}(\boldsymbol{\mu}, 1) = \ln \exp(-\mathbf{A}\boldsymbol{\mu}) = -\mathbf{A}\boldsymbol{\mu}$. This yields the common monochromatic reconstruction problem

$$\min_{\boldsymbol{\mu}\geq 0} \frac{1}{2} \|\boldsymbol{y} - \mathbf{A}\boldsymbol{\mu}\|^2 + \lambda R(\boldsymbol{\mu}), \tag{5}$$

which is treated at length by Oehler and Buzug (2007); Fessler (2000). Since projection data is preprocessed for metal artefacts and since it is based on a seemingly more robust data uncertainty model, the reconstruction algorithm derived from problem (5) provides a compelling alternative to polychromatic reconstruction. Indeed, it only requires one projection and backprojection (respectively, the product at right and at left of operator $\mathbf{A}$) per iteration, compared to two for the polychromatic algorithms.

The penalty term takes the form $R(\boldsymbol{\mu}) = \sum_{j=1}^{J} \nu_j \psi(\boldsymbol{\delta}_j^{\mathrm{t}} \boldsymbol{\mu})$, where $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2 \ldots \boldsymbol{\delta}_J$ determine image properties to penalize; ν_j weight the relative penalization of these properties and $\psi(u)$ is the penalty function. In this case, we expect the image to be of discrete structures with uniform attenuation properties. Therefore, the vectors $\boldsymbol{\delta}_j$ are rows of linear transformations of the image that determine what to penalize: here, these transformations are the identity and first-order differences. This piecewise constant assumption also prompts the use of a penalty function that preserves the discontinuities that separate the structures of the object. We use the $l_2 l_1$ penalty $\psi(u) = \sqrt{u^2 + \eta^2} - \eta$ (Charbonnier et al., 1997), which has quadratic behaviour close to the origin and linear asymptotic behaviour.

De Man et al. (2001) solves (3) with a preconditioned gradient descent method (IMPACT) that generalizes the development of Fessler et al. (1997) (who considered the maximization of the regularized Poisson log-likelihood based on a monochromatic projection model). Chueh et al. (2008) has proposed a full-image update OS-EM algorithm tailored for polychromatic reconstruction with a very similar model. Elbakri and Fessler (2002) and O'Sullivan and Benac (2007) also contributed interesting methods, which should not prove difficult to port to the projection model developed above. However, experiments by Hamelin et al. (2009*c*) on the monochromatic

reconstruction problem have shown general-purpose optimization solvers to be competitive with these EM-type methods with respect to local convergence properties. For the polychromatic reconstruction experiments performed in Sections 4 and 5, we used L-BFGS-B (Zhu et al., 1997) to solve (3), (4) and (5) directly.

Whatever iterative method is used, the runtime of the reconstruction remains excessive. Each iteration requires the computation of one or two projections and backprojections, each being costly for high-resolution image grids. The reduction of the amount of computation is addressed in the next section.

3. New efficient implementation of the polychromatic projection model

Two elements may be held responsible for the long runtimes. First, the large dimension of both the data and image vectors require at least four products (two projections and backprojections) with $\mathbf{A}$, since most L-BFGS-B iterations accept the first steplength attempt, but about one iteration in 50 backtrack once or twice). These products account for about 70% of the runtime of each iteration. Second, another large part (more than 25%) of the runtime is spent computing scalar exponentials of the modeled projections, which are summed after weighting by the emission spectrum. Both these issues are addressed through lower-dimension representations of the object (Section 3.1) and of the source spectrum (Section 3.2).

3.1. Background/ROI decomposition

For vascular imaging, high resolution is only actually needed in a small region around the blood vessel of interest. The rest of the image only serves to provide consistency to the data, so high resolution may not be necessary there.

This prompts the idea of decomposing the image into two zones: the region of interest (ROI) and the rest of the image, which we term the *background*. While good image quality is required in the former, the latter needs only be so good as to avoid hampering ROI image quality. To reduce the cost of the application of the projection operator, we propose using a coarser image grid to represent the background. This decomposition was initially proposed by Hsieh et al. (2004). The multiresolution approach to its implementation has been studied by Hamelin et al. (2009*b*) and Stearns et al. (2006) in the case of statistical reconstruction with a monochromatic projection model.

Formally, we reorder the pixels in order to decompose the object vector as $\boldsymbol{\mu} = [\boldsymbol{\mu}_{\text{bkg}}^{\text{t}}, \boldsymbol{\mu}_{\text{roi}}^{\text{t}}]^{\text{t}}$. The projection operator is similarly decomposed in the form

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{\text{bkg}}^{1} & \mathbf{A}_{\text{roi}} \\ \mathbf{A}_{\text{bkg}}^{2} & 0 \end{bmatrix},$$

where $\mathbf{A}_{\text{bkg}}^{2}$ is formed from the set of raypaths that do not intersect the ROI. The deterministic projection model then takes the form $\boldsymbol{y} = \mathbf{A}_{\text{bkg}}\boldsymbol{\mu}_{\text{bkg}} + \mathbf{A}_{\text{roi}}\boldsymbol{\mu}_{\text{roi}}$, with $\mathbf{A}_{\text{bkg}} = [(\mathbf{A}_{\text{bkg}}^{1})^{\text{t}}, (\mathbf{A}_{\text{bkg}}^{2})^{\text{t}}]^{\text{t}}$. For the polychromatic projection model (2), we need to compute the projection of both the photo-electric effect and Compton scattering contributions to the attenuation, so the model is re-expressed as

$$\tilde{\boldsymbol{y}} = \sum_{k=1}^{K} b_k \exp\left\{-\mathbf{A}_{\text{bkg}}[\Phi_k\boldsymbol{\phi}(\boldsymbol{\mu}_{\text{bkg}}) + \Theta_k\boldsymbol{\theta}(\boldsymbol{\mu}_{\text{bkg}})] - \mathbf{A}_{\text{roi}}[\Phi_k\boldsymbol{\phi}(\boldsymbol{\mu}_{\text{roi}}) + \Theta_k\boldsymbol{\theta}(\boldsymbol{\mu}_{\text{roi}})]\right\}.$$

A possible approach is the simultaneous reconstruction of the background and the ROI. However, Hamelin et al. (2007) showed that the difference in pixel size between the two images introduced a scale discrepancy that resulted in ill-conditioning of the reconstruction problem. Instead, we propose a two-steps reconstruction. In the first step, a reconstruction $\tilde{\boldsymbol{\mu}}_{\text{bkg}}$ of the full object is obtained on an image grid with the pixel size of the background. This reconstruction is expected to be rather fast, as this image grid is of low dimension. The background image $\boldsymbol{\mu}_{\text{bkg}}$ is extracted from $\tilde{\boldsymbol{\mu}}_{\text{bkg}}$ by setting to zero its pixels that belong to the ROI. In the second step, we perform a ROI-only reconstruction, using fixed background projections. With $\boldsymbol{\zeta}_{\mathbf{bkg}} \stackrel{\text{def}}{=} \mathbf{A}\boldsymbol{\phi}(\boldsymbol{\mu}_{\text{bkg}})$ and $\boldsymbol{\xi}_{\mathbf{bkg}} \stackrel{\text{def}}{=} \mathbf{A}\boldsymbol{\theta}(\boldsymbol{\mu}_{\text{bkg}})$, we devise a reconstruction algorithm around the ROI projection model

$$\tilde{\boldsymbol{y}} = \sum_{k=1}^{K} b_k \exp\left\{-\Phi_k \boldsymbol{\zeta}_{\mathbf{bkg}} - \Theta_k \boldsymbol{\xi}_{\mathbf{bkg}} - \mathbf{A}_{\text{roi}}[\Phi_k \boldsymbol{\phi}(\boldsymbol{\mu}_{\text{roi}}) + \Theta_k \boldsymbol{\theta}(\boldsymbol{\mu}_{\text{roi}})]\right\}. \tag{6}$$

This reconstruction is also fast since operator $\mathbf{A}_{\text{roi}}$ has both a lower number of rows and of columns than the original $\mathbf{A}$.

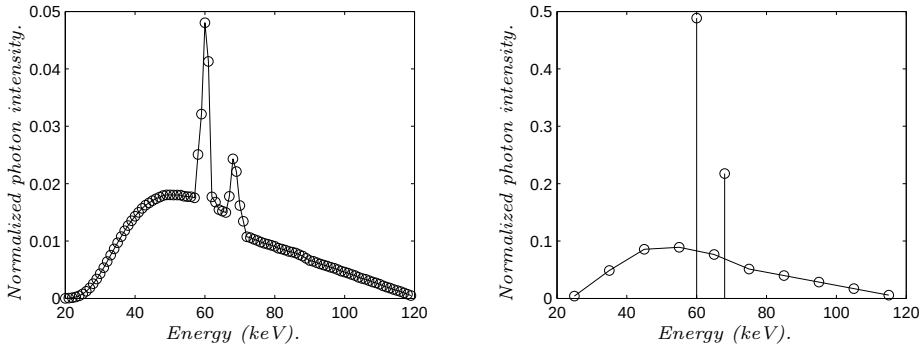
3.2. Source spectrum representation

An aspect of the polychromatic projection model that significantly inflates the runtime is the weighted summation of $K \approx 100$ sinograms. While these are generated from only two image projections, they require the computation of Kn exponentials. With the high-resolution uniform grid representation of the object, a profile of the runtime of the projection model (2) indicates that 30% of the runtime is spent on these exponentials. In addition, the X-ray emission physics is such that, on the considered energy interval, the emission spectrum is a wide-band function. A subsampling of the discrete spectrum representation suffers from significant aliasing, which hampers metal artefact reduction in the reconstruction image.

However, careful consideration of the emission phenomena can lead to a concise source spectrum model. These phenomena are general emission (*bremsstrahlung*), for which photon energy is smoothly distributed, and characteristic emission, for which generates photons of specific energy. Our approach is thus to model these phenomena separately. On the one hand, general emission admits a continuous spectrum model, which we found we could sample at a low frequency. On the other hand, since the anode from which the X photons are emitted is made of tungsten, we model characteristic emission with a pair of dirac impulses: one at 60 keV, the other at 69 keV. Let us call $b^g(E)$ the general emission spectrum and $b^c(E)$ the characteristic emission spectrum; we can then rewrite the basic projection model (1) as

$$\begin{aligned}
\tilde{\boldsymbol{y}} &= \int_0^{E_{\max}} (b^g(E) + b^c(E)) \exp[-\mathbf{A}\boldsymbol{\mu}(E)] \mathrm{d}E \\
&= \int_0^{E_{\max}} [b^g(E) + \beta_1^{\text{c}} \delta(E-60) + \beta_2^{\text{c}} \delta(E-69)] \exp[-\mathbf{A}\boldsymbol{\mu}(E)] \mathrm{d}E \\
&= \int_0^{E_{\max}} b^g(E) \exp[-\mathbf{A}\boldsymbol{\mu}(E)] \mathrm{d}E + \beta_1^{\text{c}} \exp[-\mathbf{A}\boldsymbol{\mu}(60)] + \beta_2^{\text{c}} \exp[-\mathbf{A}\boldsymbol{\mu}(69)].
\end{aligned}$$

Defining $\boldsymbol{\beta}$ as the values of a sampling of $b^g(E)$ (scaled by the sampling period) to which β_1^{c} and β_2^{c} are concatenated, we retrieve a polychromatic projection model in the summation form of (2), but with $\boldsymbol{b}$ replaced with $\boldsymbol{\beta} \in \mathbb{R}^{K'}, K' < K$.



(a) Original spectrum. (b) Remodeled spectrum (10+2).

Figure 1. Spectrum of the polychromatic source used for data simulation and its remodelling with resampling of the general emission over 10 energy values. Compared to the original spectrum, the characteristic emission impulses for the remodeled spectrum are twice as tall relatively to general emission because they cumulate the energy of the spikes. Also, the general emission contribution at the impulse energies is added only for proper graphic comparison with the original spectrum.

To compute the components of $\boldsymbol{\beta}$, we assume that for any energy level, the emitted photons either stem from general emission only or from the superposition of the two phenomena. From the full discrete representation of the spectrum, the components tainted with characteristic emission are removed. A piecewise-polynomial function is then adjusted by least squares to the remaining spectrum elements. This empirical function is taken as the continuous general emission model $b^g(E)$ and used to subtract its contribution to the spectral components composed of the two emission phenomena. What remains of these components is divided in two subsets, each respectively attributed to a characteristic emission spike. The characteristic emission energy is computed for these two spikes and the square root of this energy is used as the gain factors β_1^{c} and β_2^{c}. Finally, we have

$$\boldsymbol{\beta} = [b^g(E_1), b^g(E_2) \dots b^g(E_{K'-2}), \beta_1^{\mathrm{c}}, \beta_2^{\mathrm{c}}]^{\mathrm{t}} \in \mathbb{R}^{K'},$$

where $b^g(E_k)$ denote samples from the continuous function $b^g(E)$. An example of the remodeled source spectrum is shown in Figure 1(b).

3.3. Reconstruction problem summary

Putting these refinements together, we end up solving two polychromatic reconstruction problems: one for the full image (from which we extract the background) and one for the ROI. Both are accelerated by using the concise source spectrum model embodied by vector $\boldsymbol{\beta}$. Considering Gaussian data uncertainty, the full image reconstruction is then obtained by solving (4) on a coarse image grid, while the ROI reconstruction is obtained by solving

$$\min_{\boldsymbol{\mu}_{\mathrm{roi}} \geq 0} \frac{1}{2} \|\boldsymbol{y} - \ln \boldsymbol{p}_{\mathrm{roi}}(\boldsymbol{\mu}_{\mathrm{roi}}, \boldsymbol{\beta})\|^2 + \lambda \sum_{n=1}^{J} \nu_n \psi(\boldsymbol{\delta}_n^{\mathrm{t}} \boldsymbol{\mu}) \tag{7}$$

where

$$\boldsymbol{p}_{\mathrm{roi}}(\boldsymbol{\mu}_{\mathrm{roi}}, \boldsymbol{\beta}) = \sum_{k=1}^{K'} \beta_k \exp\left\{-\Phi_k \boldsymbol{\zeta}_{\mathbf{bkg}} - \Theta_k \boldsymbol{\xi}_{\mathbf{bkg}} - \mathbf{A}_{\mathrm{roi}} \left[\Phi_k \boldsymbol{\phi}(\boldsymbol{\mu}_{\mathrm{roi}}) + \Theta_k \boldsymbol{\theta}(\boldsymbol{\mu}_{\mathrm{roi}})\right]\right\}.$$

The reconstruction experiments of Section 5.2 validate the reconstruction algorithm derived from the Gaussian uncertainty model against the one derived from the Poisson model (3), as well as against monochromatic reconstruction (5). The former is accelerated by the background/ROI decomposition and concise source spectrum model. The latter ignores source spectrum information but may also be formulated in terms of the background/ROI decomposition, which provides excellent runtime reduction (Hamelin et al., 2009*b*). The modifications brought respectively to problems (3) and (5) for ROI reconstruction echo these that lead from (4) to (7).

4. Numerical experiments over simulated data

This first run of results aims at validating the two enhancements brought to the polychromatic projection model. The goal is to quantify the runtime reduction and ascertain the conditions under which image quality is preserved.

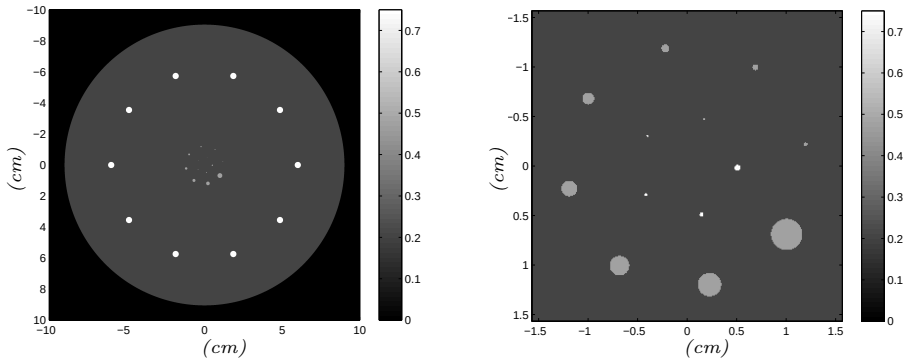
4.1. Methods

The synthetic dataset must have the same properties as one that provokes beam hardening artefacts for monochromatic algorithms. Therefore, in addition to having some metal structures, all components of the phantom must model the variation of attenuation with respect to photon energy. Moreover, the data simulation process must take into account the nonuniform emission spectrum of the X-ray source.

4.1.1. Phantom Although it is entirely virtual, the resolution phantom (shown in Figure 2) is best described in terms of physical structures and materials, because it is simulated by the generation of attenuation maps at multiple energy levels from a 2D distribution of materials. It was composed of a large ball of water inset with three concentric rings of balls. The innermost circle, mostly visible on Figure 2(b), was made of tiny inserts of iron, of varying radii. The median circle was made of larger spheres of glass, also of varying radii. The outermost circle, only visible on Figure 2(a), was made of ten larger balls of iron, all of equal radius. While the first two circles were soft and hard pixel size gauges, the outermost circle served as a perturbation to resolution measurement.

The data were simulated using the polychromatic projection model (1) with the source spectrum model represented in Figure 1(a) and the efficient projection operator code of Gendron (2008). The attenuation maps were derived from a material map of dimensions 2048×2048, in order to avoid the inverse problem crime (Wirgin, 2004). Poisson- and Gaussian-distributed pseudorandom noise were generated so as to obtain a signal-to-noise ratio (SNR) of approximately 30 dB. This hybrid noise effectively models the measurement uncertainty caused by X-ray propagation phenomena (Poisson) and electronic and hardware sources (Gaussian).

4.1.2. Selection of the regularization parameters All reconstruction problems considered above include a regularization term set up with two parameters. They are its weight λ against data adequation and η, which determines the transition between



(a) Whole field of view (20 cm). (b) Central region (3.125 cm).

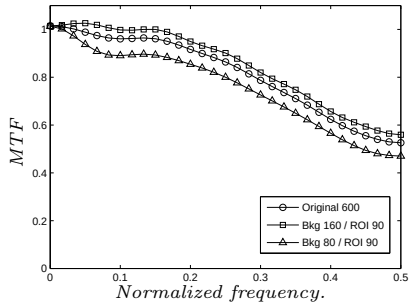
Figure 2. Resolution phantom used for data simulation (attenuation in cm^{-1} at 70 keV).

quadratic and linear behaviour of the penalty function. This is chosen heuristically as 1% of the typical discontinuity jump between different structures in the image. As this typical jump is of the order of 0.1, we set $\eta = 0.001$.

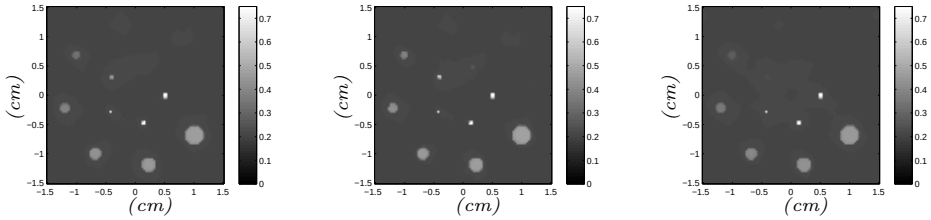
The regularization weight has a direct impact on the properties of the reconstructed image. For a low weight value, the image has high resolution and a high noise level. A higher weight value smoothes out the noise, but also the image discontinuities, resulting in a loss of resolution. We then consider two separate cases. On the one hand, the coarse full-image reconstructions are not compared to each other in terms of quality. A high weight value is thus chosen in order to reduce as much as possible de beam hardening artefacts, which dominate the reconstruction noise.

On the other hand, for images involved in quality comparison (through MTF computation as explained by Hamelin et al. (2009*c*) for simulated datasets or visual appreciation for real datasets), we try to better preserve image quality. The MTF curves for alternative reconstruction algorithms are comparable inasmuch as they are obtained from reconstructed images with a similar variance. The computation process of the MTF for simulated data yields an image variance estimate that, while not ideal, is good enough to assess MTF comparability. We then choose, using preliminary reconstructions, a weight that gives for each compared image close values of the reconstruction variance estimate (correspondance of at least the first two significant digits).

4.1.3. Nonlinear solver setup The L-BFGS-B box-constrainted nonlinear solver is used for each reconstruction problem. Memory is allocated in order to store 25 vector pairs to defined the approximative representation of the Hessian of the cost function. Also, the solver is set up to stop once the l_2 norm of the projected gradient drops below a certain tolerance. Hamelin et al. (2009*a*) justify a stopping condition based on such a stationarity measure for the truncated resolution of unconstrained linear least square problems with convex differentiable penalization. However, the set of reconstruction problems under consideration rely on a nonlinear data adequation cost. Also, the uncertainty from the ROI projection model is not accurately modeled as white Gaussian noise, on which the development of Hamelin et al. (2009*a*) is based. Still, we conjecture that stationarity is the best assertion of convergence and that the



(a) MTF.



(b) 600-pixel uniform. (c) 160/90 decomposition. (d) 80/90 decomposition.

Figure 3. Comparative reconstruction results over various background/ROI decompositions. The 160/90 decomposition essentially preserves image quality, while the 80/90 decomposition entails a sensible loss of resolution for equal variance.

results of Hamelin et al. (2009*a*) may be generalized to these experiments.

4.2. Validation of background/ROI image decomposition

The goal of this experiment is to assess the loss of image quality due to the background/ROI object representation against the reduction in runtime.

Reconstruction on two background/ROI decompositions of the object are compared to that on a high-resolution uniform image. The 3-cm ROI is represented on a 90-pixel image grid, which matches pixel size with the 600-pixel image that covers the 20-cm full field of view (FFOV). The background is extracted from either a 160-pixel or a 80-pixel image of the FFOV. All reconstructions are performed with the full source spectrum model and using the Gaussian data uncertainty distribution.

Image quality is assessed from the images and modulation transfer function (MTF) curve presented in Figure 3. The MTF curve, a 1D spectral representation of the average degradation of the image modeled as a linear space-invariant blur, is obtained by maximum-likelihood deconvolution of a finite-support point-spread function. The process is described in more details by Hamelin et al. (2009*c*). In addition, runtime for each reconstruction is displayed in Table 1.

Clearly, some background/ROI decomposition does not affect image quality. It even seems to enhance it somewhat, as the MTF of the 160/90 decomposition looks better than that of the uniform grid. Since MTF curves are highly sensitive to image degradation by smoothing, this effect may be caused by the separate calibration of the FFOV and ROI reconstructions, with the coarse-grid FFOV bearing the bulk

Table 1. Runtime of the reconstruction experiments of Section 4. Both the background/ROI object representation and concise source spectrum model decrease the runtime significantly (except for the 3+2 source model, which degenerates the polychromatic projection model).

Source	Image grid		Iterations		Runtime (min)		
	Bkg.	ROI	Bkg.	ROI	Bkg.	ROI	Total
100	- 600 -		- 991 -		—	—	338.4
100	160	90	283	160	44.9	23.4	68.3
100	80	90	164	130	25.6	27.0	52.6
100	- 600 -		- 991 -		—	—	338.4
10+2	- 600 -		- 580 -		—	—	129.0
3+2	- 600 -		- 3098 -		—	—	536.0

of artefact smoothing. Since it is also possible to vary the regularization weight spatially for the uniform-grid reconstruction, the small difference between MTF curves is inconsequential to image quality comparison.

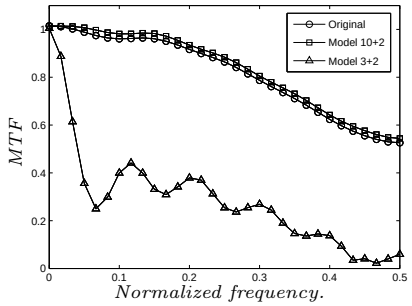
However, one can observe sharper decrease of the MTF curve for the 80/90 decomposition. This marks a limit of the background coarseness beyond which we may expect degradation of the ROI image. Although the actual degradation remains slight in this case, Hamelin et al. (2009*b*) has shown that worse results can be obtained from excessively coarse grids for the FFOV reconstruction.

Both decompositions achieve a reduction of runtime of 79% and 85% respectively, in comparison to the reconstruction on the uniform grid. Their runtimes are strongly dominated by that of the initial reconstruction, which suggests that any means of accelerating this step would translate into direct benefits.

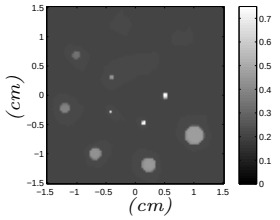
4.3. Validation of alternative source spectrum modelling

Two concise source spectrum models (as described in Section 3.2) are compared to the original discrete spectrum model in context of reconstruction. The first two models involve either sampling the general emission subspectrum on 10 elements (named the "10+2" spectrum) or 3 elements ("3+2"). All reconstructions are performed on a uniform 600-pixel image, using the Gaussian distribution of measurement uncertainty. Figure 4 presents the reconstructed image and MTF comparison. Numerical performance is displayed in Table 1.

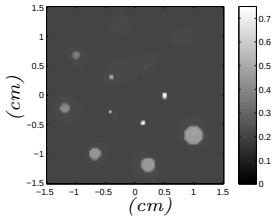
The 10+2 spectrum produces a reconstruction with the same quality as the full spectrum: the small difference that may be observed is of insignificant magnitude. This spectrum model also provides a 60% runtime decrease. Moreover, a greedy general emission subsampling such as the 3+2 spectrum yields an important loss in image quality. We may expect that as the complexity of the source spectrum model is reduced, the polychromatic projection model naturally decays to a monochromatic model. It is not surprising then that this model encurs a 61% *increase* of the runtime, as the algorithm becomes less robust to beam hardening.



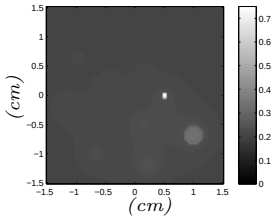
(a) MTF.



(b) Full (100) source spectrum.



(c) 10+2 spectrum.



(d) 3+2 spectrum.

Figure 4. Comparative reconstruction results over various source spectrum models. The quality of polychromatic reconstruction is preserved by the 10+2 spectrum model (Figure 1(b)), while it degrades significantly with the 3+2 model.

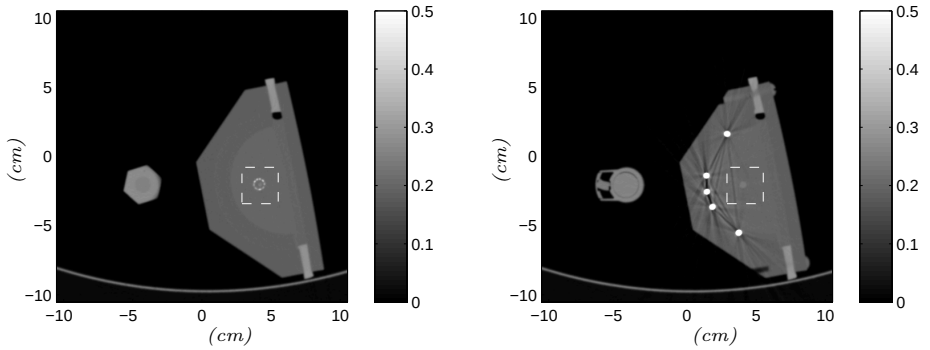
5. Reconstruction of real data

These experiments are presented as a proof of concept for polychromatic reconstruction. It is understood that the datasets studied here are preprocessed for metal artefact reduction and that monochromatic reconstruction approaches (both analytical and statistical) are invariably faster. However, an evaluation of both the feasible image quality and computational cost of polychromatic algorithms is necessary to establish their relevance.

5.1. Methods

The phantoms used for these experiments are described in Cloutier et al. (2004). They were made of one half of an hexagonal cylinder hollowed in a circular shape. This pool was filled with agar gel that mimics the attenuating properties of soft tissues. A polyurethane tube was inserted with a lost-cast material technique using sugar-based isomalt into the gel pool. A metallic stent was implanted around the tube at a middle position along the agar pool and a 70% stenosis was created in the mid portion of the stent. Metal balls were set near the bottom of the pool, on planes perpendicular to the cylinder axis. The phantom was filled with a 2.8% solution of 430 mg/ml of iothalamate meglumine (Conray 43, Mallinckrodt Medical, Pointe-Claire, Quebec, Canada) diluted with a 0.9% NaCl solution. This concentration was used to simulate the contrast obtained after bolus injection during CT angiography.

Two axial slices are of interest. The first one cuts through the stented part of the vessel model, at the site of stenosis. Since the stent is surrounding the tube,



(a) Stenosis.

(b) Metal balls in the background.

Figure 5. Scanner reconstruction of the slices of interest.

the vessel shows as a stenosed area filled with the contrast solution. The space left between the stenosed lumen and the stent struts is filled by the agar gel simulating the density of an atherosclerosis lesion. The second plane cuts through metal balls set in a non-stenosed area outside the stented portion.

The scanner used to image this phantom is a Siemens SOMATOM Sensation 16 (Siemens AG Medical Solutions, Computed Tomography, Siemensstr 1, D-91301 Forchheim, Germany) deployed at the Notre-Dame hospital of the Centre Hospitalier Universitaire de Montréal (1560 Sherbrooke Est, Montréal, Québec, Canada, H2L 4M1). The phantom was scanned using the `SpineSeq` protocol, an axial protocol similar to the helical one already deployed for contrast-enhanced vascular studies. Projections are acquired on a 672-detector array at 1160 angles around the object. The reconstruction kernel used on the scanner is geared towards noise reduction. The polychromatic X-ray source used for the scan has an angular flying focal spot, which effectively doubles the number of measurements (for a total of 1559040) and increases the image resolution that can be obtained. Figure 5 shows the images output by the scanner for both sites of interest and scan protocols used. The montage on the scanner table include two phantom assemblies buttressed against each other. The ROI is delineated with a dashed line within each image.

We compare the two polychromatic reconstruction variants to monochromatic reconstruction. The background/ROI representation of the object is used in all cases, with a 42-cm FFOV obtained on a 160-pixel image grid and a 2.625-cm ROI obtained on a 64-pixel grid. The polychromatic algorithms make use of the 10+2 source spectrum model.

Again, two regularization parameters must be chosen for each dataset and algorithm. In all cases, we set $\eta = 0.001$, for reasons described in Section 4.1.2. As for regularization weight λ, two cases are considered. For the coarse-grid FFOV reconstructions, a weight is chosen just large enough to obtain a visually satisfactory reduction of image noise. For the ROI reconstructions and for each dataset, we first determine the weight for one of the polychromatic algorithms as for the background. We then compute the empirical variance of the pixels of the image, excluding a square zone that contains the blood vessel model (and stent). We then choose the weight for the two other algorithms in order to obtain a similar image variance (excluding the vessel zone) from preliminary reconstructions.

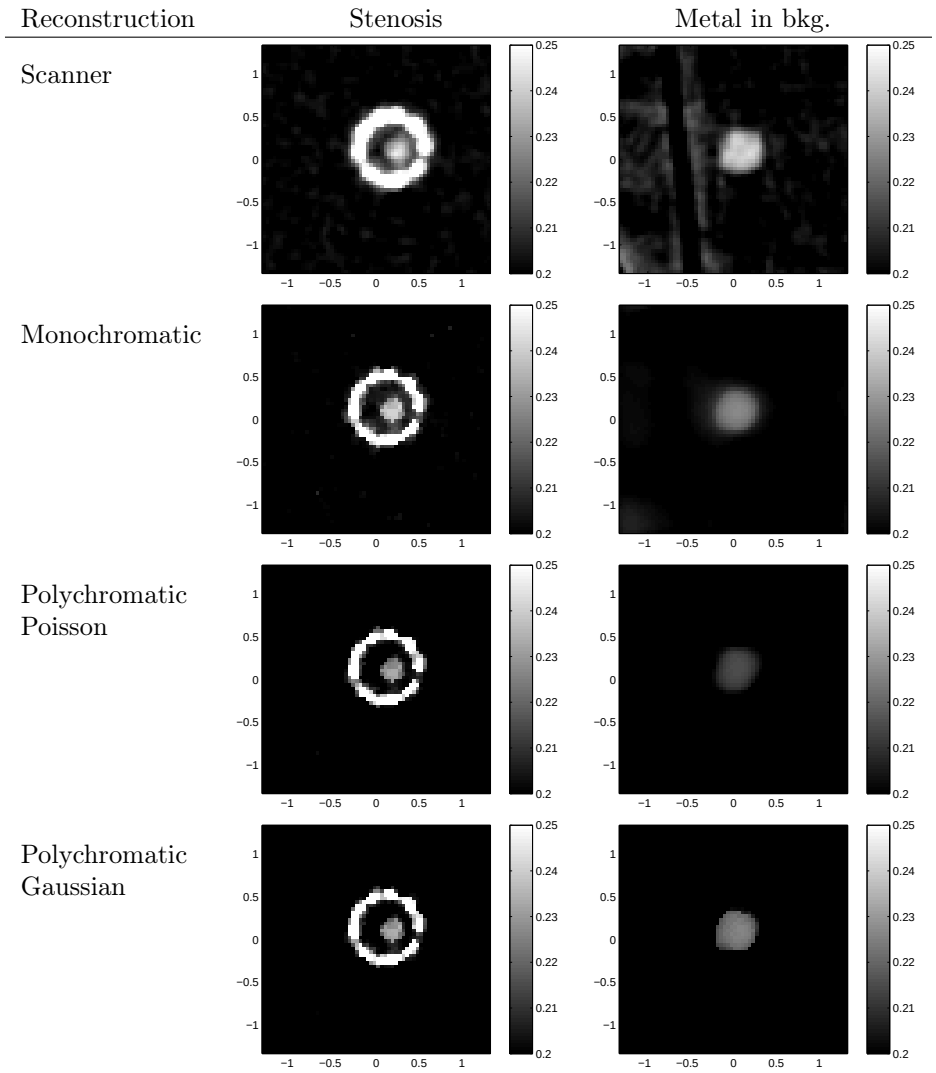


Figure 6. Real data reconstruction images. For the stenosis images, the statistical algorithms provide less noise and a better definition of stent struts. While a higher contrast between the stent struts and the intrastent lumen is obtained with the polychromatic algorithms, the vessel contour is delineated equally well for all statistical reconstructions. For the images with metal balls in the background, the scanner reconstruction suffers from strong artefacts, which are practically eliminated for all statistical reconstructions. The Gaussian polychromatic variant yields the best resolution.

5.2. Results

Figure 6 shows the reconstruction results for the stenosis-and-stent and metal-balls-in-background slices. The ROI from the scanner reconstructions is displayed along the results for the statistical reconstructions algorithms, for reference. All images are rescaled to map the gray levels on a scale from 0.2 to 0.25. The runtime for each reconstruction case is reported in Table 2.

Table 2. Comparative results of the reconstruction of real datasets. The monochromatic algorithm is the fastest in all cases. The Gaussian polychromatic variant arrives second in all cases, converging in between 2 and 3 times as much time. The Poisson polychromatic variant converges after 5 to 7 times as much.

Slice	Projection model	Iterations		Runtime (min)		
		Bkg.	ROI	Bkg.	ROI	Total
Stent	Mono.	133	36	5.4	0.1	5.5
	Poly. Poisson	287	25	24.4	0.5	24.9
	Poly. Gauss.	146	31	15.2	0.9	16.1
Balls	Mono.	160	30	7.1	0.1	7.2
	Poly. Poisson	572	21	45.8	0.6	46.4
	Poly. Gauss.	191	35	16.4	0.7	17.1

We observe that the statistical algorithms suppress most artefacts present in the scanner images. Let us first consider the stenosis images. It may be seen that stent wire thickness is less exaggerated and that the contour of the blood vessel is better contrasted to the background. This spread of the stent wires in all cases likely stems from data preprocessing in order to reduce the metal artefacts, which has succeeded since no streak nor shadow can be observed, even on the scanner image. The best contrast of the stent struts to the intrastent lumen is obtained for the Gaussian polychromatic variant. Still, the contour of the vessel is rendered with enough contrast for all statistical methods.

In the second slice, where the background contains metal balls, data preprocessing could not completely compensate for the low photon counts that yield the beam hardening artefacts, as the scanner images show. The statistical methods then provide a significant reduction of these remaining artefacts. However, for the monochromatic and Poisson polychromatic algorithms, heavier regularization weights had to be set to reduce the artefacts, resulting in poor resolution. Such regularization settings must be chosen when the data doesn't match the projection model well according to the log-likelihood being minimized. For the monochromatic algorithm, poor data adequation may be traced down to the improper modelling of the X-ray beam used for imaging. For the Poisson polychromatic variant, the Poisson log-likelihood embodies a good modelling of the uncertainties induced by the physics of X-ray propagation and detection. Therefore, if the dataset is modified for the purposes of reducing artefacts of various sources (see Section 1), the uncertainty distribution will shift away from the Poisson probabilistic model.

For all considered datasets, the monochromatic algorithm has come out the fastest by a large margin and the Gaussian polychromatic variant is second. This poor numerical performance of the Poisson polychromatic variant is also a symptom of poor adequation of the projection model to data with respect to the log-likelihood. It appears that the Poisson polychromatic algorithm lacks robustness against data preprocessing. This contrasts with the Gaussian variant, which seems to be less sensitive to preprocessing-induced data inconsistency in all respects. It is likely that for a raw dataset, the Poisson and Gaussian variants perform similarly, both in terms of image quality and runtime.

Finally, for objects containing only small metal objects (such as stent wires), the metal artefact reduction realized through sinogram preprocessing appears sufficient.

Therefore, the polychromatic algorithms do not yield significantly better images than their monochromatic counterpart, which converges much faster. However, the artefacts induced by larger metal structures cannot be sufficiently eliminated by the preprocessing, in which case the Gaussian polychromatic algorithm proposed in this paper should be preferred.

6. Conclusion

We have presented multiple refinements to a polychromatic projection model and reconstruction algorithm based of the work of De Man et al. (2001). These refinements yield a runtime reduction by one order of magnitude. However, this is not sufficient to recommend the algorithm for vascular imaging, where only small metal structures such as a stent are encountered. Monochromatic reconstruction is rather recommended in this case, for its sheer speed and its good resolution-to-noise compromise.

Future work will focus on an extension to 3D statistical reconstruction, for both monochromatic and polychromatic projection models. The main challenge is that even for coarse axial grids, the runtime of statistical FFOV reconstruction is inconvenient. We would then like to extract the background from the scanner reconstructions.

Acknowledgments

This work was funded by Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Collaborative Health Research Project #323572-06 of the Canadian Institutes of Health Research. B. Hamelin also acknowledges a Postgraduate Scholarship from the NSERC.

References

Alvarez R E and Macovski A 1976 *Phys. Med. Bio.* **21**(5), 733–744.

Charbonnier P, Blanc-Féraud L, Aubert G and Barlaud M 1997 *IEEE Trans. Image Processing* **IP-6**(2), 298–311.

Chueh H S, Tsai W K, Chang C C, Chang S M, Su K H and Chen J C 2008 *Comput. Meth Programs Biomed.* **92**, 289–293.

Cloutier G, Soulez G, Qanadli S D, Teppaz P, Allard L, Qin Z, Cloutier F and Durand L G 2004 *Med. Phys.* **31**(6), 1424–1433.

De Man B, Nuyts J, Dupont P and Suetens P 2001 *IEEE Trans. Medical Imaging* **20**(10), 999–1008.

Elbakri I A and Fessler J A 2002 *IEEE Trans. Medical Imaging* **21**(2), 89–99.

Elbakri I A and Fessler J A 2003 *Phys. Med. Bio.* **48**(7), 2453–2477.

Fessler J A 1994 *IEEE Trans. Medical Imaging* **13**(2), 290–300.

Fessler J A 2000 *in* J. M Fritzpatrick and M Sonka, eds, 'Handbook of Medical Imaging' Vol. 2 SPIE Press Bellingham, WA chapter 1, pp. 1–70.

Fessler J A, Ficaro E P, Clinthorne N H and Lange K 1997 *IEEE Trans. Medical Imaging* **16**(2), 166–175.

Gendron D 2008 Méthodes algébriques pour la reconstruction en tomographie à rayons X Master's thesis École Polytechnique de Montréal.

Hamelin B, Goussard Y and Dussault J P 2007 *in* '29th Annual International Conference of the IEEE EMBS' pp. 739–742.

Hamelin B, Goussard Y and Dussault J P 2009*a* On the stopping criterion for numerical methods used to solve linear systems with additive Gaussian noise Technical Report EPM-RT-2009-10 École Polytechnique de Montréal. `http://www.optimization-online.org/DB_HTML/2009/09/2404.html`.

Hamelin B, Goussard Y, Dussault J P, Cloutier G, Beaudoin G and Soulez G 2009*b* Submitted to *Med. Phys.* .

Hamelin B, Goussard Y, Dussault J P, Cloutier G, Beaudoin G and Soulez G 2009*c* Submitted to *IEEE Trans. Medical Imaging* .

Hamelin B, Goussard Y, Gendron D, Dussault J P, Cloutier G, Beaudoin G and Soulez G 2008 *in* 'Proc. IEEE ISBI' pp. 1453–1456.

Hsieh J, Thibault J B, Bouman C and Sauer K 2004 'Iterative method for region-of-interest reconstruction' US Patent.

Kak A C and Slaney M 1987 *Principles of Computerized Tomographic Imaging* IEEE Press New York, NY.

LaRivière P J, Bian J and Vargas P A 2006 *IEEE Trans. Medical Imaging* **25**(8), 1022–1036.

LaRivière P J and Vargas P 2006 *in* M. J Flynn and J Hsieh, eds, 'Medical Imaging 2006: Physics of Medical Imaging' Vol. 6142 Proc. SPIE pp. 61420X–1–61420X8.

Mou X, Tang S, Luo T, Zhang Y and Yu H 2008 *in* J Hsieh and E Samei, eds, 'Medical Imaging 2008: Physics of Medical Imaging' Vol. 6913 Proc. SPIE pp. 691333–1–10.

Müller J and Buzug T M 2009 *in* E Samei and J Hsieh, eds, 'Medical Imaging 2009: Physics of Medical Imaging' Vol. 7258 Proc. SPIE pp. 72581Y–1–8.

Oehler M and Buzug T M 2007 *Meth. Inform. Med.* **46**, 261–269.

Oehler M, Kratz B, Knopp T, Müller J and Buzug T M 2008 *in* P. J Bones, M. A Fiddy and R. P Millane, eds, 'Image Reconstruction from Incomplete Data V' Vol. 7076 Proc. SPIE San Diego pp. 707607–1–10.

O'Sullivan J A and Benac J 2007 *IEEE Trans. Medical Imaging* **26**(3), 283–297.

Rinkel J, Dillon W P, Funk T, Gould R and Prevrhal S 2008 *J. Comput. Assist. Tomogr.* **32**(4), 621–629.

Stearns C W, Majeshwar R M and Wollenweber S D 2006 *in* 'IEEE Nucl. Sci. Symp.' pp. 2808–2811.

Wirgin A 2004 'The inverse crime'.
URL: *http://arxiv.org/abs/math-ph/0401050v1*

Zhu C, Byrd R H, Lu P and Nocédal J 1997 *ACM Trans. Math. Soft.* **23**(4), 550–560.