

## Iterative correction of beam hardening artifacts in CT.

G. Van Gompel<sup>1,2</sup>, K. Van Slambrouck<sup>3</sup>, M. Defrise<sup>4</sup>, K. J. Batenburg<sup>1,5</sup>, J. de Mey<sup>2</sup>, J. Sijbers<sup>1</sup>, and J. Nuyts<sup>3</sup>

<sup>1</sup> *IBBT-Vision Lab, Universiteit Antwerpen, Antwerp, Belgium*

<sup>2</sup> *Department of Radiology, UZ Brussel - VUB, Brussels, Belgium*

<sup>3</sup> *Department of Nuclear Medicine, K.U. Leuven, Leuven, Belgium*

<sup>4</sup> *Department of Nuclear Medicine, Vrije Universiteit Brussel, Brussels, Belgium*

<sup>5</sup> *Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands*

**Purpose.** To reduce beam hardening artifacts in CT in case of an unknown X-ray spectrum and unknown material properties.

**Methods.** We assume that the object can be segmented into a few materials with different attenuation coefficients, and parameterize the spectrum using a small number of energy bins. The corresponding unknown spectrum parameters and material attenuation values are estimated by minimizing the difference between the measured sinogram data and a simulated polychromatic sinogram. Three iterative algorithms are derived from this approach: two reconstruction algorithms IGR and IFR, and one sinogram pre-correction method ISP.

**Results.** The methods are applied on real X-ray data of a high and a low-contrast phantom. All three methods successfully reduce the cupping artifacts caused by the beam polychromaticity in such a way that the reconstruction of each homogeneous region is to good accuracy homogeneous, even in case the segmentation of the preliminary reconstruction image is poor. In addition, the results show that the three methods tolerate relatively large variations in uniformity within the segments.

**Conclusions.** We show that even without prior knowledge about materials or spectrum, effective beam hardening correction can be obtained.

## I. INTRODUCTION

10 When a monochromatic X-ray beam traverses a homogeneous object, the total attenuation coefficient is linearly related to the thickness of the object along that ray (Beer's law). In general, however, CT X-ray sources are polychromatic. The linear relation does not hold for polychromatic beams, since lower energy photons are more easily absorbed than higher energy photons, which causes the beam to 'harden' as it propagates through the  
15 object. This non-linear effect is referred to as beam-hardening. If the energy dependence of the absorption is not taken into account, reconstructions are contaminated by cupping and streak artifacts<sup>1</sup>. Beam hardening correction is important in both medical and industrial CT applications to improve the visual quality of the images and to obtain more accurate segmentations, which is necessary for morphometric image analysis.

20 Beam hardening artifacts have been a subject of research for decades, resulting in a broad variety of artifact reduction strategies. Beam hardening correction methods can roughly be subdivided into four classes : hardware filtering, dual energy, statistical polychromatic reconstruction, and linearization.

25 Hardware filtering is a common method to narrow the broad source spectrum. Thin metal plates that are placed between the source and the object absorb the lower energy photons of the beam before the beam enters the object. Although hardware filters reduce the beam hardening artifacts in the resulting image, the lower photon count also results in a decrease of the signal to noise ratio.

30 In dual energy methods<sup>1-3</sup>, the energy-dependency of the attenuation coefficients is modeled as a linear combination of two basis functions representing the separate contributions of the photo-electric effect and the scattering. The coefficients of the two basis functions are needed for each image pixel. Therefore, two scans at different source voltages are required, preferably with non-overlapping spectra<sup>3</sup>. After determining the coefficients, reconstructions  
35 of the linear attenuation coefficient can be estimated at any energy within the diagnostic range. Dual energy CT in medical imaging is typically limited to specific applications e.g. for measuring the bone mineral density in the lumbar spine<sup>4</sup>, improved detection of lung embolism<sup>5</sup>, and plaque detection in coronary arteries using fast kV switching<sup>6</sup>.

Statistical reconstruction of polychromatic data is an approach explored by several

40 authors<sup>7-10</sup>. The statistical beam hardening reduction methods basically incorporate the polychromatic nature of the beam in a maximum likelihood (ML) algorithm. This approach assumes that the object consists of  $N$  known base substances, and that the energy dependence of the attenuation coefficient for each pixel can be described as a linear combination of the known energy dependencies of the base substances. Statistical methods are very  
45 flexible with respect to various geometries, prior knowledge, noise statistics, etc. However, such methods are computationally very expensive.

Linearization methods aim to transform the measured polychromatic attenuation data into monochromatic attenuation data. For homogeneous objects, the correction is typically computed using the beam hardening curve, which describes the attenuation-thickness relation of the material, and which is acquired from a calibration scan<sup>11,12</sup>. For objects containing  
50 more than one material, this method can be extended<sup>13</sup>. Alternatively, the linearization for multiple material objects can be performed using an iterative post reconstruction (IPR) approach<sup>1,12,14-22</sup>. IPR methods are initialized with a preliminary reconstruction of the data, which is used to estimate the intersection length of each material with each ray path.  
55 Using these material thicknesses and prior knowledge about the spectrum and materials, the sinogram can be corrected. The resulting reconstruction image can then be used to initialize a new iteration yielding improved estimates of the material thicknesses and consequently improved beam hardening correction. Similarly to the statistical methods, the linearization techniques assume that the object consists of a known number of materials with known  
60 energy dependence of the attenuation coefficients. Some methods require uniform materials (e.g.<sup>16,18</sup>), others allow for mixtures of these base materials (e.g.<sup>1,19-21</sup>).

An important limitation of the statistical and linearization methods is the prior knowledge about the materials and the energy spectrum of the source-detector system that is  
65 required. In many industrial and in some medical cases (e.g. prostheses), the exact material composition is unknown, hence, the required prior knowledge is not available.

A few linearization methods have been developed that do not require prior knowledge such as the spectrum and the attenuation coefficients of the materials. These methods  
70 only require the number of materials  $N$  to be known and assume that the object consists of uniform materials. The methods of Van de Castele et al.<sup>23</sup>, Gao et al.<sup>24</sup> and Mou et

al.<sup>25</sup> offer nice results but have some limitations. The first two methods are limited to a particular class of objects, while the third method is restricted to complete data and uses a cost function based on data consistencies, which is specific for the used 2D fan beam  
75 geometry.

Recently, Krumm et al.<sup>26</sup> proposed an iterative method similar to IPR method but which does not require prior knowledge on the spectrum or material attenuations. The method assumes  $N$  materials and segments an initial reconstruction, thus obtaining for each ray  
80 an intersection length with each material. These  $N$  values correspond to a point in a  $N + 1$ -dimensional hyperspace. A smooth hypersurface and an hyperplane are fitted to the set of points obtained from all measured rays, representing respectively a polychromatic and monochromatic approximation to the attenuation values. The difference between the two surfaces is computed for each measured ray and used as an additive correction. For mild  
85 beam hardening problems, i.e. when an adequate segmentation can be performed based on the uncorrected reconstruction, the method of Krumm et al. successfully suppresses beam hardening artifacts in the reconstruction of piecewise constant objects.

In this paper, we propose an alternative approach based on a physical model. It is as-  
90 sumed that the spectrum can be represented with a small and predefined number of energy bins, and that the object consists of a known number of different materials. We propose iterative procedures that 1) segment the object, assigning each pixel to one of the materials, 2) estimate the amplitude of each energy bin and 3) estimate the attenuation coefficient of each material for each of the energy bins. The parameter estimation is treated as an  
95 optimization problem, minimizing the difference between the observed measurements and the polychromatic projection computed from the estimated parameters. The algorithms do not require the materials to be perfectly uniform, they can deal with small density variations within each material.

Based on this approach, three algorithms are proposed. The iterative gradient based  
100 reconstruction (IGR) minimizes the cost function by alternately updating a subset of the parameters. To allow for non-uniformities within each material, a relative density is estimated for each pixel. This relative density estimation is the most expensive part. With an iterative filtered backprojection approach (IFR), we have attempted to accelerate this

step by replacing an adjoint operation (backprojection) with a pseudo inverse operation  
 105 (filtered backprojection). IGR and IFR are reconstruction algorithms, which update an im-  
 age by minimizing the difference between the measurement and a polychromatic simulation.  
 Applying similar assumptions in the sinogram domain and omitting the relative density  
 parameters leads to an iterative sinogram preprocessing method (ISP), where a sinogram is  
 updated based on the difference between a monochromatic and polychromatic simulation,  
 110 as proposed by<sup>14,26,27</sup>.

The proposed algorithms require the number of materials to be known but do not use  
 any information about the source-detector spectrum or material properties. The price to  
 pay is that the reconstructed values cannot be interpreted as accurate estimates of the at-  
 115 tenuation coefficients at a specific energy because the algorithms do not attempt to estimate  
 the actual energy corresponding to each energy bin. The goal, however, is to correct the  
 cupping and streaking artifacts caused by beam hardening.

The paper is outlined as follows. Section 2 introduces the notations and concepts, and  
 120 Section 3 describes the polychromatic model. The IGR, IFR and ISP methods are described  
 in Section 4. The proposed methods have been evaluated for real X-ray CT data of 2- and  
 3-material hardware phantoms, and in addition on simulated clinical data, as described in  
 Sections 5 and 6. Finally, the conclusion is found in Section 7.

## II. NOTATIONS AND CONCEPTS

125 Assume an object that consists of  $N$  materials with attenuation coefficients  $\mu_n(E)$ ,  $n =$   
 $1 \dots N$ , that depend on the energy  $E$ . Denote by  $J$  the number of pixels on the reconstruction  
 grid. We assume that each pixel contains exactly one material defined by a vector of binary  
 variables  $\mathbf{s} = \{s_{n,j}\}$ , with  $s_{n,j} = 1$  if voxel  $j$  contains material  $n$ , and  $s_{n,j} = 0$  otherwise.  
 Consider projection lines, denoted by an index  $i = 1 \dots D$ , for which an intensity measure is  
 130 obtained at the detector, and denote by  $l_{i,j}$  the intersection length of ray  $i$  with pixel  $j$ . We  
 also introduce the relative density  $d_j$  of the material in each pixel, with  $\mathbf{d} = \{d_j\}$ , to allow  
 the reconstruction of small non-uniformities within each segment. When a monochromatic  
 X-ray beam with intensity  $I_0$  and energy  $E_0$  passes through the object along a path  $i$ , the

monochromatic exit intensity for that specific ray can be expressed as (Beer-Lambert).

$$I_{\text{mono},i} = I_0 e^{-\sum_{n=1}^N \mu_n(E_0) \sum_{j=1}^J l_{i,j} d_j s_{n,j}}. \quad (1)$$

We assume that  $I_0$  is independent of  $i$  for convenience. If an air scan is available, the possible position dependence of  $I_0$  can easily be taken into account.

The monochromatic attenuation  $A_{\text{mono}}$  for a given energy  $E_0$  is defined as the logarithm of the ratio of the input and output intensities for that energy, and behaves linearly with respect to the traversed thickness:

$$A_{\text{mono},i} = \sum_{n=1}^N \mu_n(E_0) \sum_{j=1}^J l_{i,j} d_j s_{n,j}. \quad (2)$$

Note that  $A_{\text{mono},i}$  is the desired quantity used by analytic reconstruction algorithms. We use FBP in the text and equations to denote any analytic reconstruction algorithm suitable for the given 2D or 3D geometry, e.g. filtered backprojection (FBP), FDK<sup>28</sup>, helical cone beam reconstruction<sup>29</sup>.

In practice, the emitted X-ray photons have varying energies  $E \in [0, E_{\text{max}}]$  and also the detector response is energy dependent. The measured intensity of such a polychromatic beam along a path  $i$  can be expressed as the sum of the monochromatic contributions for each energy  $E$ :

$$I_{\text{poly},i}^{\text{meas}} = \int_0^{E_{\text{max}}} I_0(E) e^{-\sum_{n=1}^N \mu_n(E) \sum_{j=1}^J l_{i,j} d_j s_{n,j}} dE, \quad (3)$$

where  $I_0(E) = I_{\text{source}}(E)\epsilon(E)$  is referred to below as the ‘‘source-detector energy spectrum’’, equal to the product of the incident energy spectrum  $I_{\text{source}}(E)$  with the detector efficiency  $\epsilon(E)$ .

The polychromatic attenuation  $A_{\text{poly},i}^{\text{meas}}$  along a path  $i$  that is used by non-corrected algorithms is defined by

$$A_{\text{poly},i}^{\text{meas}} = \log \left( \frac{I_0}{I_{\text{poly},i}^{\text{meas}}} \right), \quad (4)$$

<sup>135</sup> with  $I_0$  the total incident beam intensity  $I_0 = \int_0^{E_0} I_0(E) dE$ .

### III. POLYCHROMATIC MODEL

Assume that the attenuation coefficients  $\mu_n(E)$  and the source spectrum  $I_0(E)$  are unknown. We model the source-detector energy spectrum as a discrete set of  $E_m$  energy bins with intensities  $\{I_e\}$  ( $e = 1 \dots E_m$ ). Using this model, the output polychromatic intensity  $I_{\text{poly},i}^{\text{sim}}$  for a path  $i$  is given by

$$I_{\text{poly},i}^{\text{sim}} = \sum_{e=1}^{E_m} I_e e^{-\sum_{n=1}^N \mu_{n,e} \sum_{j=1}^J l_{i,j} d_j s_{n,j}}, \quad (5)$$

where  $\boldsymbol{\mu} = \{\mu_{n,e}\}$  represents a set of effective attenuation coefficients corresponding to each material/energy bin pair. With this model, the polychromatic output attenuation  $A_{\text{poly}}^{\text{sim}}$  is computed using

$$A_{\text{poly},i}^{\text{sim}} = -\log \left( \sum_{e=1}^{E_m} I_e^F e^{-\sum_{n=1}^N \mu_{n,e} \sum_{j=1}^J l_{i,j} d_j s_{n,j}} \right), \quad (6)$$

where  $I_e^F = \frac{I_e}{\sum_e I_e}$  represents the fraction of the total spectrum intensity corresponding to the  $e^{\text{th}}$  energy bin.

In our approach, the aim is to find the reconstructed image for which the difference between the corresponding simulated polychromatic sinogram  $A_{\text{poly},i}^{\text{sim}}$  and the measured sinogram  $A_{\text{poly},i}^{\text{meas}}$  is minimized in a least square sense. In other words, the aim is to minimize the cost function  $\Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{d}, \mathbf{s})$  defined by

$$\Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{d}, \mathbf{s}) = \frac{1}{D} \sum_{i=1}^D \left( \log \left( \frac{I_{\text{poly},i}^{\text{meas}}}{I_0} \right) - \log \left( \sum_{e=1}^{E_m} I_e^F e^{-\sum_{n=1}^N \mu_{n,e} \sum_{j=1}^J l_{i,j} d_j s_{n,j}} \right) \right)^2, \quad (7)$$

where the attenuation coefficients  $\boldsymbol{\mu} = \{\mu_{n,e}\}$ , relative bin intensities  $\mathbf{I}^F = \{I_e^F\}$ , the segmentation  $\mathbf{s} = \{s_{n,j}\}$  and the relative densities  $\mathbf{d} = \{d_j\}$  are the parameters to be estimated.

<sup>140</sup> Note that the cost function defined in Eq.(7) is not convex, and since  $\Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{d}, \mathbf{s}) = \Phi(\boldsymbol{\mu}/2, \mathbf{I}^F, 2\mathbf{d}, \mathbf{s})$  its minimization clearly does not have a unique solution. The three methods proposed below aim at finding an interesting (local) minimum. It is therefore important to initialize the optimization procedures with a good estimate of the parameters.

## IV. METHODS

### 145 A. Gradient based optimization (IGR)

The IGR approach defines the problem as an explicit optimization problem. Each operation in the iteration attempts to decrease the cost function (Eq. (7)), and is guaranteed not to increase it.

150 The IGR method is initialized by performing a preliminary reconstruction  $R^0 = \text{FBP}(A_{\text{poly}}^{\text{meas}})$  and selecting the number of materials  $N$  and of energy bins  $E_m$ . Each iteration  $w$  of the algorithm consists of the following steps:

#### 1. Determine parameters $\mathbf{s}$

Segment the image  $R^{w-1}$  into  $N$  segments by thresholding. In the first iteration ( $w = 1$ ), the segmentation is performed by k-means clustering<sup>30</sup>. For  $w > 1$ , the thresholds are adapted in such a way that the new thresholds minimize the cost function (Eq. (7)),

$$\mathbf{s}^w = \arg \min_{\mathbf{s}} \Phi(\boldsymbol{\mu}^{w-1}, \mathbf{I}^{F,w-1}, \mathbf{d}^{w-1}, \mathbf{s}) \quad (8)$$

where the notation  $\arg \min_{\mathbf{s}}$  stands for the optimization of the cost function over the set of segmentations  $\mathbf{s}$  that can be generated by varying the threshold parameters. If no set of thresholds is found that decreases  $\Phi$ , the current segmentation is used without  
 155 update in the next iteration, so  $\mathbf{s}^w = \mathbf{s}^{w-1}$ .

#### 2. Update of the relative density

In the first iteration ( $w = 1$ ), the relative densities are assigned the value  $d_j = 1$  for  $j = 1 \dots J$ . For  $w > 1$ , the relative density parameters  $\mathbf{d}^w$  are determined using a gradient descent algorithm (see Appendix A) such that

$$\mathbf{d}^w = \arg \min_{\mathbf{d}_{\geq 0}} \Phi(\boldsymbol{\mu}^{w-1}, \mathbf{I}^{F,w-1}, \mathbf{d}, \mathbf{s}^w). \quad (9)$$

#### 3. Attenuation and fractional intensity

Compute updated attenuation coefficients  $\boldsymbol{\mu}^w$  and fractional intensities  $\mathbf{I}^{F,w}$  by minimizing the cost function  $\Phi$  at constant relative density and segmentation, using a gradient based algorithm:

$$\left( \boldsymbol{\mu}^w, \mathbf{I}^{F,w} \right) = \arg \min_{\boldsymbol{\mu} > 0, \mathbf{I}^F \geq 0} \Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{d}^w, \mathbf{s}^w), \quad (10)$$



with the constraint  $\mu_{n,e_1} > \mu_{n,e_2}$  if  $e_1 < e_2$ . In the first iteration, the seed values of the intensity fractions for the optimization procedure are set at  $I_j^F = 1/E_m$ . The optimization of  $\mu_{n,e}$  in the first iteration is initialized by computing the mean attenuation  $\langle \mu \rangle_n$  in each of the segments and assigning the initial values  $0.2\langle \mu \rangle_n$ ,  $\langle \mu \rangle_n$ , and  $5\langle \mu \rangle_n$  to the three energy bins.

#### 4. Image update

At this point all reconstruction parameters  $(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{d}, \mathbf{s})$  are updated. A new image is needed for segmentation and display purposes only. We arbitrarily choose the following updated image  $R_j^w$  for the segmentation in the next iteration:

$$R_j^w = d_j^w \sum_n \bar{\mu}_{IGR,n}^w s_{n,j}^w. \quad (11)$$

where  $\bar{\mu}_{IGR,n}^w$  is the median of  $\{\mu_{n,e}^w, e = 1 \dots E_m\}$ .

Concerning the practical implementation, a few remarks are in order:

- Recall that, since the cost function is non-convex, only a local minimum can be found. The notation *arg min* should be interpreted as reaching a local minimum.
- For the relative density update step, the algorithm described in appendix A is implemented using data subsets. Only one pass through the complete dataset is performed and consequently only a lowering of the cost function is obtained.
- To derive a stop criterion for the IGR algorithm, we define the polychromatic model error  $\epsilon^w$  of iteration  $w$  to be the value of the cost function  $\Phi(\boldsymbol{\mu}^w, \mathbf{I}^{F,w}, \mathbf{d}^w, \mathbf{s}^w)$ . This polychromatic model error is easily monitored, and is guaranteed to decrease monotonically as the number of iterations increases. The iterative procedure is terminated once following criterion is met:

$$\frac{\epsilon^w + \epsilon^{w-1}}{\epsilon^{w-2} + \epsilon^{w-3}} > t, \quad (12)$$

with  $0 < t < 1$  a threshold value. The averaging between pairs of successive iterations aims at improving the robustness of the stop criterion.

## B. FBP-based optimization (IFR)

An important disadvantage of the proposed IGR method is its computational complexity. The determination of the relative densities  $d_j^w$  (step 2) represents the main computational bottleneck because several backprojection operations need to be performed. In the IFR method we attempt to accelerate this step by replacing the back projections with a pseudo inverse operation: filtered backprojection.

The proposed FBP-based optimization method (IFR) is identical to the IGR method, except for the relative density update step (2), which is replaced by

$$\mathbf{d}^w = \mathbf{d}^{w-1} + \boldsymbol{\omega}^w \cdot \text{FBP}(A_p^{\text{meas}} - A_p^{\text{sim}}(\boldsymbol{\mu}^{w-1}, \mathbf{I}^{F,w-1}, \mathbf{d}^{w-1}, \mathbf{s}^w)), \quad (13)$$

where the dot denotes an elementwise multiplication and  $\boldsymbol{\omega}$  is a diagonal relaxation factor with elements

$$\omega_j^w = \frac{1}{\sum_{n=1}^N \mu_{\text{IFR},n}^{w-1} s_{n,j}^w} \quad (14)$$

where we chose  $\mu_{\text{IFR},n}^w = \max\{\mu_{n,e}^w, e = 1, \dots, E_m\}$  since it yields a conservative step size. This relaxation factor is empirical (see also “Link with the IFR method” in IV C) but ensures that the update step is correct from the dimensional point of view.

The iteration (13) reaches a fixed point if the simulated polychromatic sinogram exactly matches the measured data. However, the choice of the relaxation factor (14) is empirical and therefore, in contrast with the IGR method, there is no guarantee that the update step (13) decreases the value of the cost function, though this was observed in the numerical experiments in section 6. The IFR method proceeds with the same steps (3) and (4) as the IGR method.

## C. Sinogram Preprocessing method (ISP)

The large number of relative density parameters to be estimated in the IGR and IFR methods might in some cases negatively influence the conditioning of the optimization problem and the stability in the presence of noise. Though standard regularization techniques might be used to alleviate this problem, we propose in this section a third method (ISP), which does not require relative density parameters (i.e.  $d_j = 1$  for all pixels) while still allowing to recover structures that have been inaccurately segmented.

Starting with an initial image estimate  $R^0 = \text{FBP}(A_{\text{poly}}^{\text{meas}})$ , each ISP iteration consists of  
 195 the following steps:

### 1. Determine parameters $\mathbf{s}$

The parameters  $\mathbf{s}^w$  are determined by thresholding the image  $R^{w-1}$ . The thresholds are adapted as for the IGR and IFR method.

### 2. Attenuation and fractional energy

Updated attenuation coefficients and fractional energies are calculated as in IGR and IFR:

$$\left(\boldsymbol{\mu}^w, \mathbf{I}^{F,w}\right) = \arg \min_{\boldsymbol{\mu}, \mathbf{I}^F} \Phi(\boldsymbol{\mu}, \mathbf{I}^F, \mathbf{d} = \mathbf{1}, \mathbf{s}^w). \quad (15)$$

except that all relative densities are set equal to 1.

### 3. Reference attenuation and mono- and polychromatic simulation

Reference attenuation coefficients  $\bar{\mu}_{\text{ISP},n}^w$  are determined (see below for the derivation of the reference attenuation coefficients). The monochromatic and polychromatic simulations are then calculated using respectively

$$A_{\text{mono},i}^{\text{sim},w} = \sum_{n=1}^N \bar{\mu}_{\text{ISP},n}^w \sum_{j=1}^J l_{i,j} s_{n,j}^w. \quad (16)$$

and eq. (6) with  $d_j^w = 1$ ,

$$A_{\text{poly},i}^{\text{sim},w} = -\log \left( \sum_{e=1}^{E_m} I_e^{F,w} e^{-\sum_{n=1}^N \mu_{n,e}^w \sum_{j=1}^J l_{i,j} s_{n,j}^w} \right), \quad (17)$$

### 4. Corrected sinogram and image update

A monochromatized sinogram  $A^{\text{corr},w}$  is calculated using

$$A^{\text{corr},w} = A_{\text{poly}}^{\text{meas}} + (A_{\text{mono}}^{\text{sim},w} - A_{\text{poly}}^{\text{sim},w}) \quad (18)$$

200 and the next image estimate is reconstructed as  $R^w = \text{FBP}(A^{\text{corr},w})$ .

As for the previous methods, the iterations are terminated when the stop criterion (eq. (12)) is satisfied.

The ISP method is similar to sinogram pre-processing methods of the type proposed by Nalcioglu and Lou<sup>27</sup>, which monochromatize the polychromatic measurement by adding a  
 205 correction  $C$ , given in this case by  $C = A_{\text{mono}}^{\text{sim},w} - A_{\text{poly}}^{\text{sim},w}$ . In the two first methods (IGR

and IFR) described above, the reconstructed image is given by equation (11), and small deviations from piecewise constant images as well as segmentation errors could be recovered only thanks to the presence of the relative density parameters. The reason why the ISP method allows modeling such small deviations despite the fact that relative densities are fixed as  $d_j^w = 1$  for all  $w$  and all pixels  $j$  is that the corrected sinogram (18) still contains the original measurements.

### *Reference attenuation coefficients*

To derive a set of reference attenuation coefficients for sinogram correction, note that the updated sinogram in Eq. (18) is composed of the original measured sinogram and a correction term based on the estimated parameters. Assume a case in which the segmentation procedure classifies two low-contrast materials as one material. Since the relative density parameters, which compensate for misclassification in the IGR and IFR method, are set to  $d_j = 1$  for all  $j$  in the ISP method, information about the discrimination of the materials is only present in the measured data, not in the correction term. We expect that low contrast features will be optimally visible in the updated reconstruction image if the reference attenuation coefficients  $\bar{\boldsymbol{\mu}}_{\text{ISP}} = \{\bar{\mu}_{\text{ISP},n}\}$  used in Eq. (16) minimize the magnitude of the correction term in Eq. (18), i.e. if they minimize the following quadratic functional with respect to  $\bar{\boldsymbol{\mu}}$ :

$$\begin{aligned} \Psi(\bar{\boldsymbol{\mu}}_{\text{ISP}}^w) &= \sum_i^D (A_{\text{mono},i}^{\text{sim},w} - A_{\text{poly},i}^{\text{sim},w})^2 \\ &= \sum_i^D \left( \sum_{n=1}^N \bar{\mu}_{\text{ISP},n}^w t_{n,i}^w + \log \left( \sum_{e=1}^{E_m} I_e^{F,w} e^{-\sum_{n=1}^N \mu_{n,e}^w t_{n,i}^w} \right) \right)^2. \end{aligned} \quad (19)$$

with

$$t_{n,i}^w = \sum_{j=1}^J l_{i,j} s_{n,j}^w. \quad (20)$$

To solve this minimization problem, define a matrix  $\mathbf{B}^w$  with elements

$$b_{n,n'}^w = \sum_i^D t_{n,i}^w t_{n',i}^w \quad n, n' = 1 \dots N, \quad (21)$$

and a vector  $\mathbf{v}$  as

$$v_n^w = - \sum_i^D t_{n,i}^w \log \left( \sum_{e=1}^{E_m} I_e^{F,w} e^{-\sum_{n=1}^N \mu_{n,e}^w t_{n,i}^w} \right) \quad n = 1 \dots N. \quad (22)$$

The minimum of the functional in Eq. (19) is then given by

$$\bar{\boldsymbol{\mu}}^w = \mathbf{B}^{w+} \mathbf{v}, \quad (23)$$

225 where  $\mathbf{B}^{w+}$  denotes the Moore-Penrose generalized inverse of matrix  $\mathbf{B}^w$ .

#### *Link with the IFR method*

There is a link between the IFR and ISP methods. Assuming that FBP is the exact inverse of the system matrix defined by the  $l_{i,j}$ , and multiplying the IFR update step by  $1/\omega^w$ , one  
230 finds that (13) is equivalent to

$$\sum_j^J \sum_{n=1}^N \mu_{\text{IFR},n}^{w-1} s_{n,j}^w l_{i,j} d_j^w = A_{\text{poly}}^{\text{meas}} + \sum_j^J \sum_{n=1}^N \mu_{\text{IFR},n}^{w-1} s_{n,j}^w l_{i,j} d_j^{w-1} - A_{\text{poly}}^{\text{sim}}(\boldsymbol{\mu}^{w-1}, \mathbf{I}^{F,w-1}, \mathbf{d}^{w-1}, \mathbf{s}^w) \quad (24)$$

Identifying the LHS with  $A^{\text{corr},w}$  and the second term in the RHS with  $A_{\text{mono}}^{\text{sim},w}$  in (18), we see that the two methods are closely related, although the identification is of course not exact since all relative densities are set equal to 1 in the ISP method.

## D. Acceleration

235 When applying iterative beam hardening methods as presented in this paper to mild beam hardening problems, the number of iterations is very small. In Section V, examples of challenging beam hardening artifacts are presented for which the number of iterations that is required to meet condition (12), is quite large. Performing a large number of reconstructions on a whole 3D dataset is not desirable. We address this problem by performing a preliminary  
240 fast iterative beam hardening correction on a downsampled sinogram and image grid. In addition, we found that the number of iterations can further be reduced by smoothing the reconstructed image prior to the segmentation in each iteration.

We implemented the following strategy to accelerate the three algorithms:

Stage (1) Apply the beam hardening reduction method on a downsampled sinogram until the  
245 stop criterion (12) is reached. In each iteration, a Gaussian filter is used to smooth the updated image before segmentation.

Stage (2) Use the resulting sinogram (for the ISP method) or reconstruction image (IGR and IFR) of step (1) to initialize a new series of beam hardening correction iterations without Gaussian smoothing, until condition C is met.

Stage (3) Upsample the corrected sinogram/image resulting from step (2) to the size of the original measured sinogram. Finalize the algorithm by performing iterations until the stop criterion (12) is reached.

In addition we used ordered subsets for projections and backprojections in the relative density update step of IGR (see Eq. (A11))<sup>31</sup>.

## 255 V. EXPERIMENTS

The proposed beam hardening artifact reduction methods are evaluated using two dedicated physical phantoms: the *Bean* and the *Barbapapa* phantom (see Fig. 1). Both phantoms consist of a non-convex polymethylmethacrylate (PMMA) body in which several holes are drilled. The three holes of the low-contrast Bean phantom are filled with air, mineral spirit  
260 and water, the latter having a low physical contrast with PMMA. The Barbapapa phantom has five holes of which three are filled with aluminum, representing a high-contrast problem.

X-ray data of both phantoms were acquired using a SkyScan 1172  $\mu$ CT scanner with circular cone beam geometry. The tube voltage was deliberately set as low as 60 kV to  
265 induce challenging beam hardening artifacts, and hardware filtering and the software beam hardening correction option were turned off. We selected only the central slice and rebinned the data to a parallel beam sinogram, consisting of 300 equally spaced views with angular range  $[0, \pi[$ , and 1000 radial samples. The images were reconstructed on a  $1000 \times 1000$  grid with a pixel size of  $25.16 \mu m$ .

270

To simulate medical data, a digital segment image was generated based on a real CT  
275 scan of a human body at shoulder height. The segment image, shown in Fig. 1(c) consists of two uniform materials (adipose tissue and soft tissue), and one material (cortical bone) with

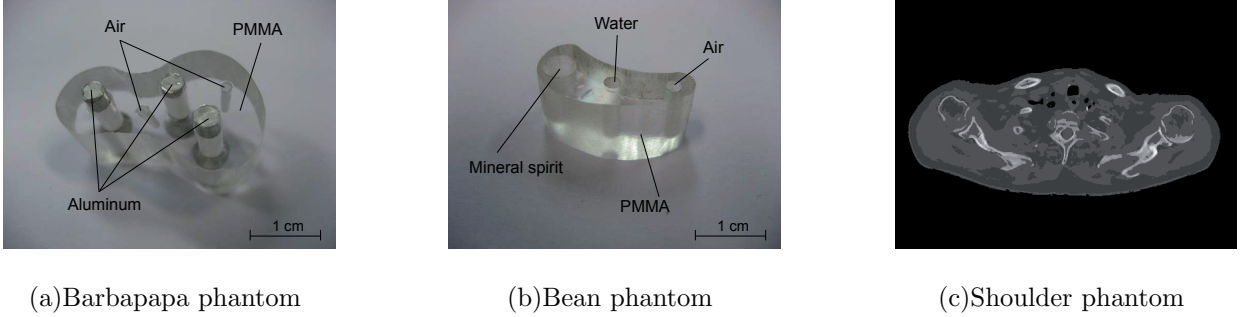


Figure 1. Pictures of the phantoms. (a) and (b) are hardware phantoms; (c) is a simulated phantom

non-uniform density ( $(\text{density}_{\max} - \text{density}_{\min})/\text{density}_{\min} = 2$ ). Realistic polychromatic data was simulated using a tungsten spectrum of 60 kV, discretized in 30 bins of 2 keV, obtained from Report 78<sup>32</sup>. No energy-dependent detector response was taken into account. The corresponding attenuation values for each material were obtained from the ICRU-44 report<sup>33</sup>. The simulated polychromatic data was then obtained by forward projection using Eq. (4). No scatter or inhomogeneous bow-tie filtering were considered.

To correct beam hardening artifacts for the three phantoms, we followed the acceleration strategy described in Section IV D with the stop criterion parameter  $t = 0.97$ . The methods were initialized using downsampled sinograms and a Gaussian smoothing filter with a standard deviation of 1 pixel. For the Barbapapa and Bean phantom, the downsampled sinograms contained 250 radial by 150 angular samples, yielding reconstruction images on a  $250 \times 250$  pixels grid with pixel size  $100.64\mu\text{m}$ . The downsampled sinogram of the Shoulder phantom contained 256 radial by 180 angular samples.

All beam hardening correction experiments were performed using 3 energy bins and  $N = 3$  materials (including air).

## VI. RESULTS AND DISCUSSION

### A. Beam hardening correction for Barbapapa and Bean phantom

Fig. 2 shows the FBP reconstruction of the Barbapapa phantom from the uncorrected measured sinogram. In this image, important cupping artifacts can be observed in the

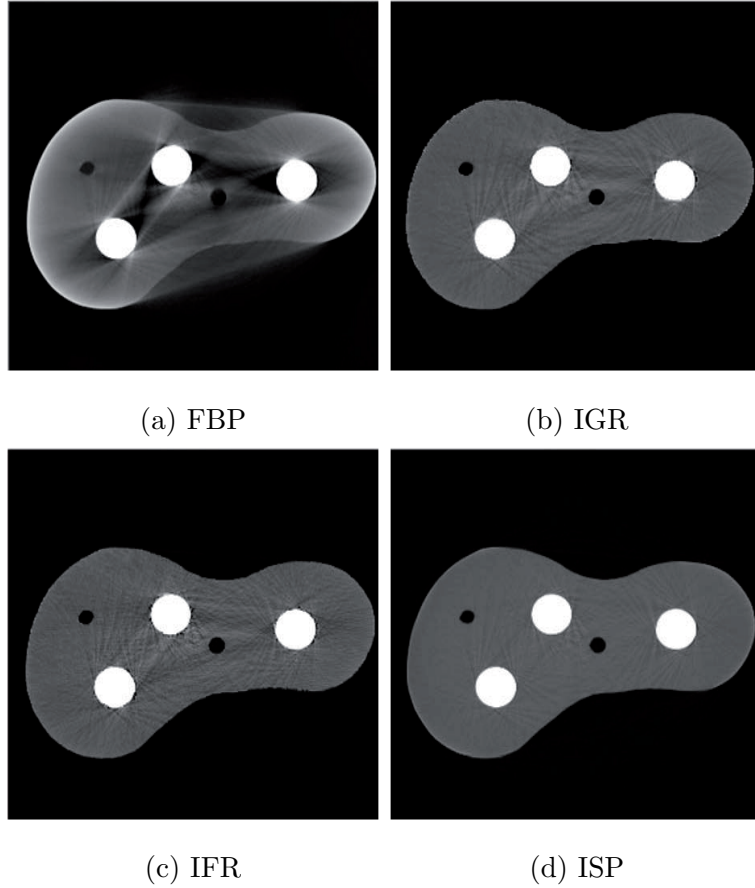


Figure 2. Reconstructions from experimental data of the Barbapapa phantom.

PMMA and aluminum area. In addition, the background attenuation in the convex hull of the phantom is overestimated. Note also the streak artifacts in between the aluminum rods,  
 300 which could mask air holes.

The IGR, IFR and ISP methods were evaluated for the Barbapapa phantom. The resulting reconstruction images are depicted in Fig. 2(b), (c) and (d), respectively. The IGR and IFR images are displayed using the monochromatic  $\bar{\mu}_{\text{IGR},n}^w$ . From these images and the  
 305 corresponding line profiles along the central horizontal line, which are plotted in Fig. 3 (a), (b) and (c), it can be observed that all three methods strongly suppress the cupping and streak artifacts compared to a non-corrected FBP reconstruction. Note that the meaning of the reconstructed image values is different for IGR and IFR compared to ISP. IGR and IFR produce three images with attenuation coefficients, one for each energy bin. In contrast,  
 310 ISP produces a single sinogram, which after reconstruction yields a single image with



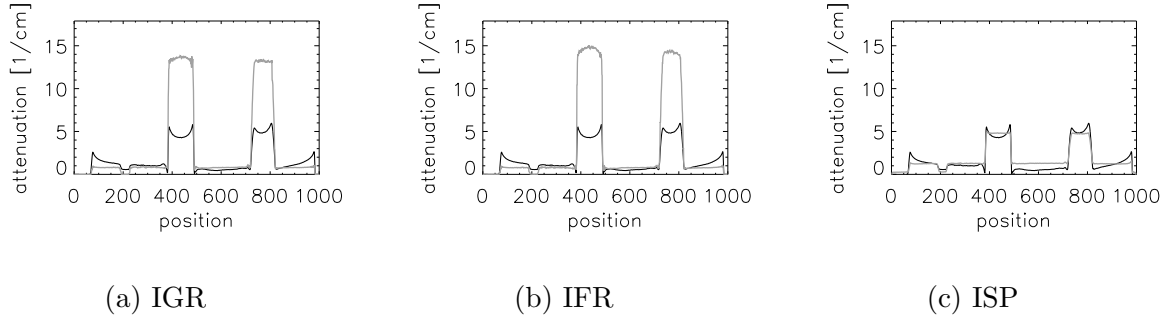


Figure 3. Line profiles along the vertically centered horizontal line in the reconstruction image of the Barbapapa phantom. The black line corresponds to the uncorrected FBP reconstruction; the grey line represents the corrected reconstruction.

effective attenuation coefficients. ISP also yields  $N \times E_m$  estimates for the attenuation of each material at each energy bin, but these are only used during the computation of the sinogram updates. The relation between these estimates and the reconstructed attenuation values is complex and depends on the object and on the energy spectrum.

315

Fig. 4 shows results for the Bean phantom. Fig. 4(a), (b), (d) and (e) correspond to the uncorrected FBP, the IGR, IFR and ISP reconstructions from the 60 kV unfiltered scan, respectively. The line profiles corresponding to the latter three methods, along the line shown in Fig. 5 (a), are depicted in Fig. 5 (b) to (d). Again, a strong suppression of the cupping artifacts is observed for all three methods.

320

The Bean phantom challenges the contrast sensitivity of the methods. Due to the low CT value contrast between water and PMMA, those materials are considered as one material by the segmentation procedure. However, in all three reconstruction images, the water area can be discriminated from the surrounding PMMA. For the IGR and IFR methods, this is due to the introduction of the relative density parameter, which compensates the assignment of the water pixels to PMMA by a slight increase in relative density. In the ISP method, on the other hand, the polychromatic model will not differentiate between the water and PMMA. However, by maximizing the contribution of the measured data to the corrected sinogram through the selection of the reference attenuation coefficients (see section IV C), the ISP method does have a mechanism to tolerate non-uniform features in

330

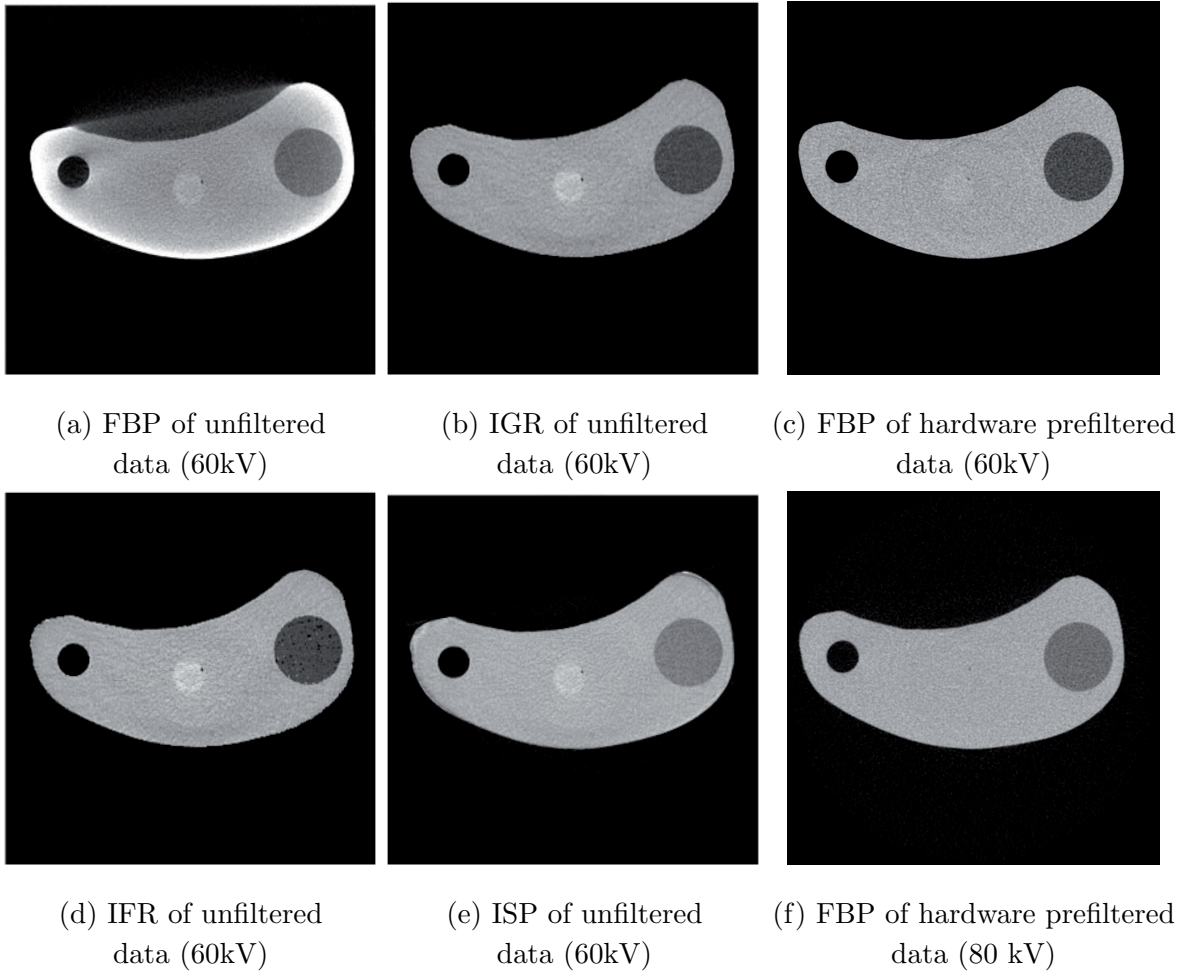


Figure 4. Reconstructions from experimental data of the Bean phantom.

the images. It can be concluded that, despite the assumption of uniform segments, the methods tolerate a certain degree of non-uniformity.

335

As a comparison, we also show the reconstruction image from X-ray data obtained with hardware prefiltering of the X-ray beam. Fig. 4(c) and (f) depict the standard FBP reconstruction from X-ray data of the same Bean phantom acquired with Al-Cu prefiltering of a 60kV and a 80kV beam, respectively, whereas figures (a), (b), (d) and (e) were obtained from unfiltered 60kV data. The filtered data were acquired using a higher tube current as to compensate for the reduced incident flux on the object caused by prefiltering. The images show that beam prefiltering effectively reduces beam hardening artifacts, but also reduces the contrast between water and PMMA, due to the increased effective energy of the entrance beam.

340

345

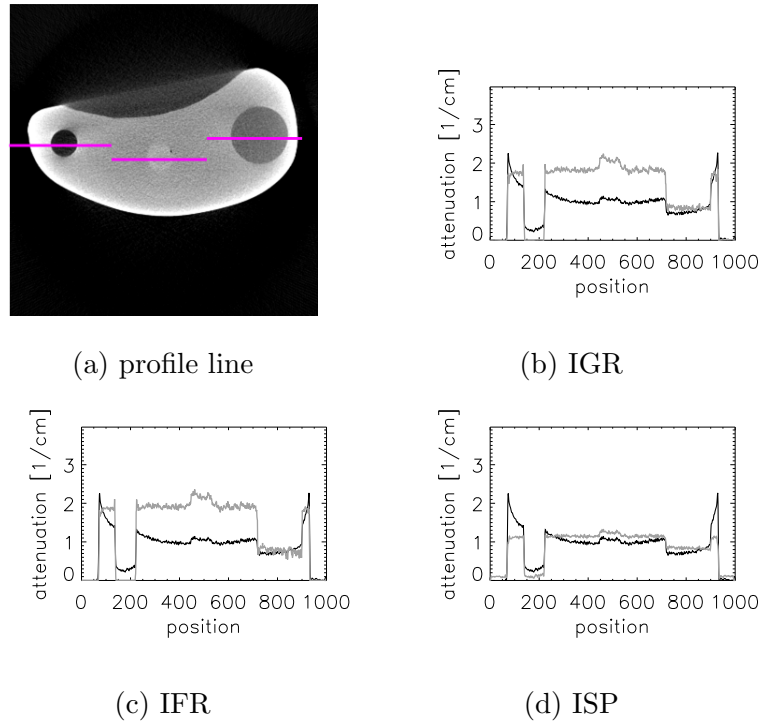


Figure 5. Profiles in the reconstruction images of the IGR, IFR and ISP methods along the broken profile line shown in (a) for the Bean phantom. The black line corresponds to the uncorrected FBP reconstruction; the grey line represents the corrected reconstruction.

Fig.6 and Fig.7 show the results of the beam hardening correction for the Shoulder phantom, displayed at a low-contrast enhancing grey scale and a large grey scale range, respectively. To show the stability of the algorithms for larger and structurally complicated non-homogeneous areas, the beam hardening correction was performed using three materials (N=3) corresponding to air, soft+adipose tissue, and bone (with varying relative densities). The line profiles corresponding to the IGR, IFR and ISP methods are depicted in Fig. 8 (a) to (c). It can be observed from Fig. 6, Fig. 7 and Fig. 8 that, although the material segments are clearly non-homogeneous, the cupping artifacts are removed and an important reduction of streaks between bone parts is obtained. These results suggest that the methods might be applicable for medical or preclinical data.

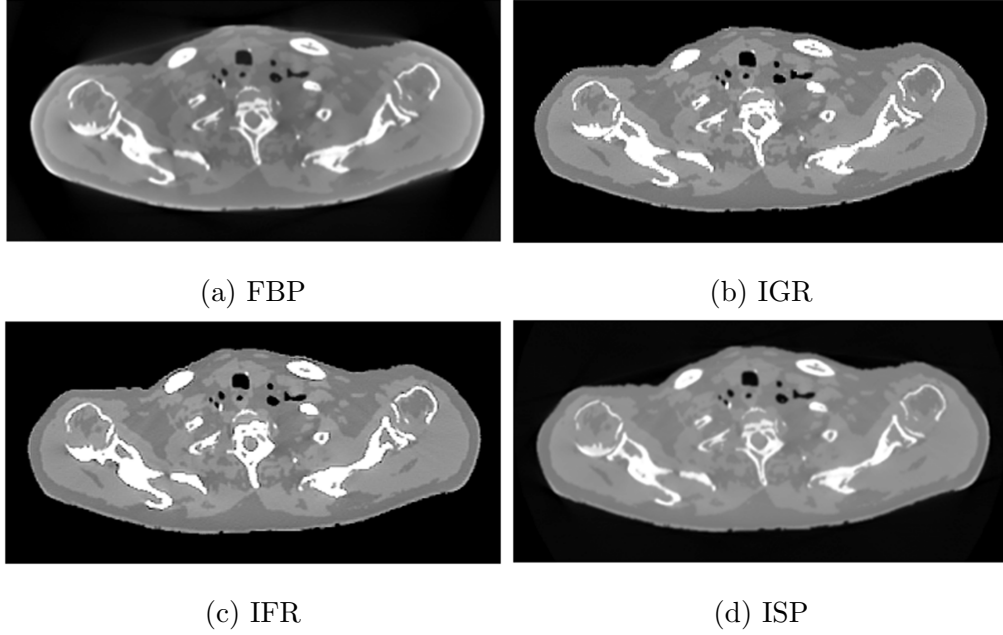


Figure 6. Reconstructions from experimental data of the Shoulder phantom (soft tissue low contrast enhancing gray scale range).

## B. Convergence

The cost function that is optimized in each iteration is non-convex, and therefore none of the three algorithms can guarantee convergence to a global optimum. In practice, we found that the obtained parameter estimations were dependent of the initial seed parameters, but  
 360 without noticeable impact on the reconstruction images.

In Fig. 9 the evolution of the model error throughout successive iterations is plotted for the Barbapapa (a) and the Bean phantom (b). The IGR, IFR and ISP methods used in total 76, 47 and 17 iterations, respectively, to run through the successive acceleration steps of the scheme in Section IV D and reach the final stop criterion for the Barbapapa phantom.  
 365 For the Bean phantom, these numbers are respectively 45, 39 and 17, and for the Shoulder phantom 97, 45 and 19 iterations. The peaks in Fig. 9 correspond to transitions to the next acceleration stage. Although only the IGR method guarantees a monotonic decrease of the cost function value as a function of the iteration number within a single acceleration stage, the IFR and ISP method show a similar trend. The model error converges more slowly in  
 370 the IGR method compared to the IFR method, which can be explained by the nature of the optimization. In each iteration, IFR estimates the relative density map  $d_i$  using the FBP

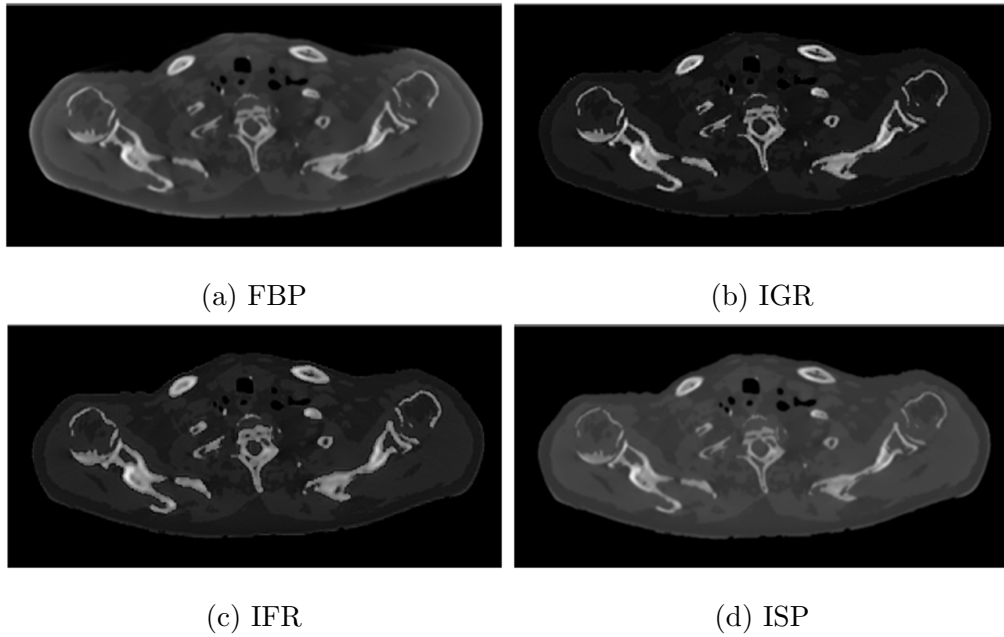


Figure 7. Reconstructions from experimental data of the Shoulder phantom (large grey scale range).

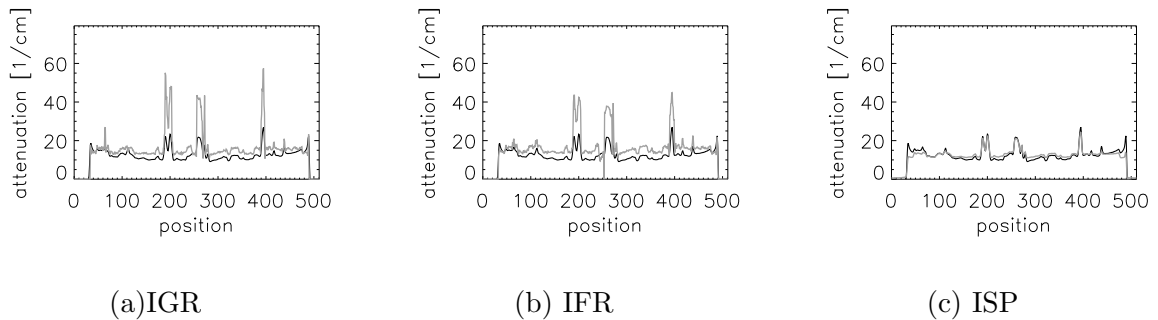


Figure 8. Line profiles along the vertically centered horizontal line in the reconstruction image of the Shoulder phantom. The black line corresponds to the uncorrected FBP reconstruction; the grey line represents the corrected reconstruction.

in which the filtering enhances the high frequencies. In the IGR method, where the relative density map is computed using a limited number of least squares optimization steps, the high frequencies are introduced more slowly.

375 A shared observation for both phantoms for the IGR and IFR method is the similar value of the model error, obtained after convergence. This could be expected since IGR and IFR minimize the same cost function. For the ISP method, on the other hand, a higher model

error is obtained as a consequence of the much lower number of degrees of freedom in the cost function. This, however, does not imply an inferior image quality compared to the IFR and IGR images, since the ISP method has an alternative mechanism in sinogram space to tolerate non-uniformities within segments.

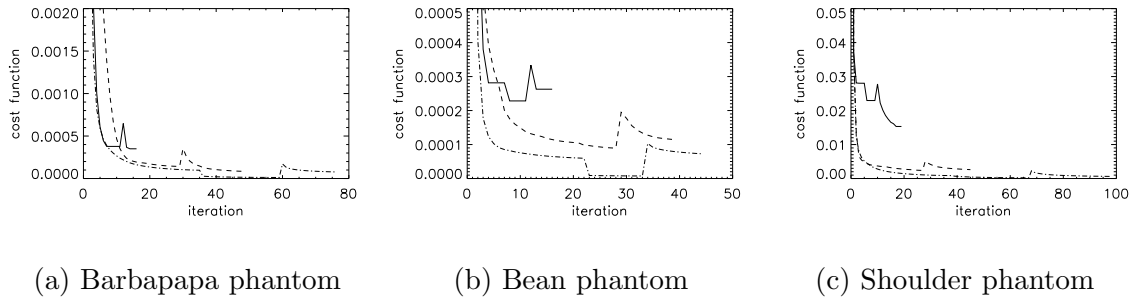


Figure 9. Evolution of the cost value as a function of iteration number for the IGR, IFR and ISP methods. The dash-dotted, dashed and solid curves denote the cost function for the IGR, IFR and ISP methods, respectively.

### C. Segmentation

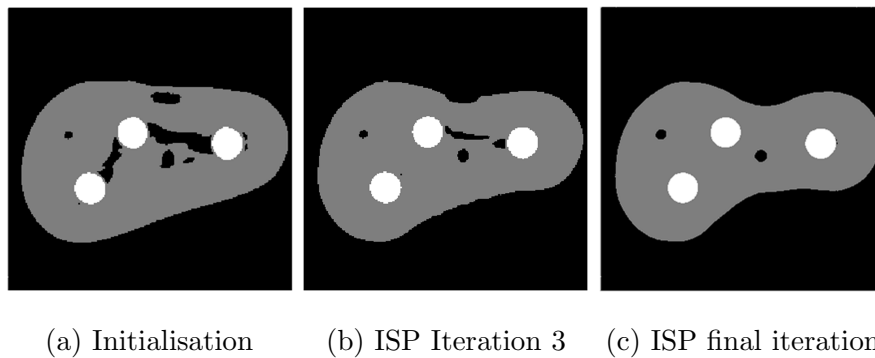


Figure 10. Illustration of the segmentation improvement in successive iterations of the beam hardening correction procedure.

Fig. 10(a) shows the segmentation of the first FBP reconstruction. Note that quantitative measurements based on this segmentation would introduce potentially important errors. The major challenge for the reconstruction of this phantom is to obtain a correct segmentation from this primary image. The center and right images of Figure 10 show the gradual

improvement of the segmentation image through successive iterations for the ISP method.  
 390 A similar trend (not shown) was observed for the IGR and IFR method. The images demonstrate that the three methods are able to recover accurate reconstruction images with realistic segmentations, even when the segmentation of the initial reconstruction is poor. As a segmentation with global thresholding is suboptimal, a significant reduction of the number of iterations might be obtained using sophisticated segmentation strategies, e.g.  
 395 adaptive thresholding by projection distance minimization<sup>34</sup>. The selection of appropriate segmentation methods is beyond the scope of this paper.

#### D. Computational complexity

The major factor affecting the computational complexity of the three methods is the number of back- and forward projections (both  $O(P^2Q)$ ) with P and Q respectively the number  
 400 of detector pixels and angles, as shown in Table I. Recall that iterative postreconstruction (IPR) beam hardening correction methods, which were described in the Introduction, require only one back- and one forward projection per iteration, whereas the proposed algorithms described in Section IV, require in each iteration multiple backprojections for the segmentation, and the IGR requires in addition multiple back- an forward projections for the relative  
 405 density update. Note, however, that other segmentation strategies such as histogram based thresholding could be considered, which would then reduce the computational complexity of IFR and ISP to one back- and one forward projection per iteration, like IPR methods.

	IGR	IFR	ISP
Segmentation	N per segmentation iteration	N per segmentation iteration	N per segmentation iteration
Relative density update	N+2	1	not applicable
Attenuation and fractional intensities	0	0	0
Image Update	not applicable	not applicable	1

Table I. Number of back- or forward projections used in the different steps of the IGR, IFR and ISP algorithms (see Section IV); N is the number of materials.

## E. Interpretation of the image and the estimated parameters

An approximate analysis of the reconstructed attenuation coefficients for the Barbapapa phantom is given in Fig. 11, using the source-detector energy spectrum and material characteristics which were unknown to the algorithms. For each energy bin and each material, the algorithms yield an intensity  $I_e^F$  and an attenuation coefficient  $\mu_{n,e}$  with dimension  $cm^{-1}$ . Using the material properties for aluminum and PMMA, these estimated attenuation coefficients were used to assign energies  $\tilde{E}_{n,e}$  to each energy bin and each material, and these assignments turned out to be uniquely determined in this case. The results are plotted in Fig. 11 as a set of impulses of height  $I_e^F$  positioned at energies  $\tilde{E}_{n,e}$  and superimposed on top of the real spectrum. Although the number of materials and energy bins is too small to draw definite conclusions, the figure suggests that the obtained set of discrete energy peaks is compatible with the source-detector spectrum. For a (nearly) monochromatic spectrum, the methods are expected to yield (almost) exact attenuation coefficients corresponding to the spectrum energy.

Fig. 12 presents an alternative analysis of the same data, which compares the attenuation coefficients obtained by the three methods with the known material attenuation curves (solid lines) of aluminum (grey) and PMMA (black). The energy boundaries in the figure are obtained by dividing the spectrum of Fig. 11 in three bins, such that the integrals over the bins correspond to the estimated fractions  $I_e^F$ , with  $e = 1..E_m$ . The estimated attenuation coefficients  $\mu_{n,e}$ , indicated with stars and plotted in the center of the corresponding energy bins, are within or close to the  $\mu$  intervals determined by the known attenuation curves for each energy bin.

## VII. CONCLUSION

In this paper, we described three iterative methods for beam hardening correction for objects consisting of multiple, uniform materials. The methods are implemented for a 2D parallel beam geometry, but are readily extendable to any acquisition geometry. The three methods aim at minimizing a cost function based on a discretized model for polychromatic attenuation. The IGR method, and its accelerated approximating method



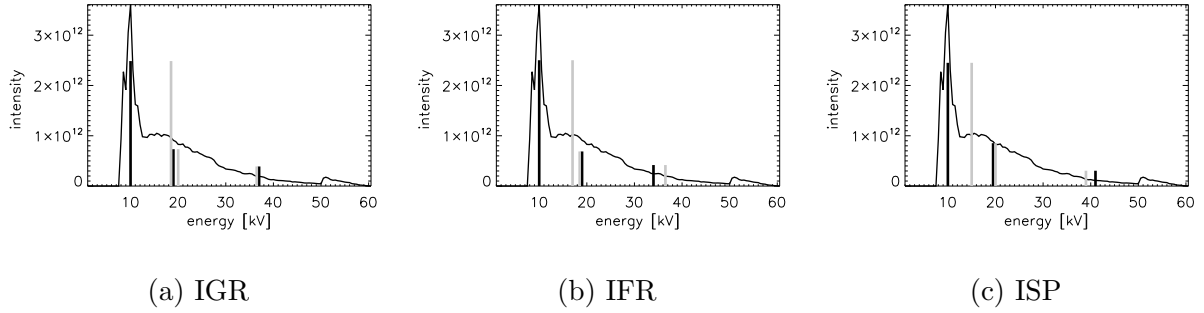


Figure 11. Comparison for the Barbapapa phantom of the real spectrum with the discrete spectrum that is composed (see text) from the obtained set of attenuation coefficients and corresponding fractional intensities, using also the known material properties (that were unknown to the methods). The grey and black lines represent the results for aluminum and PMMA, respectively.

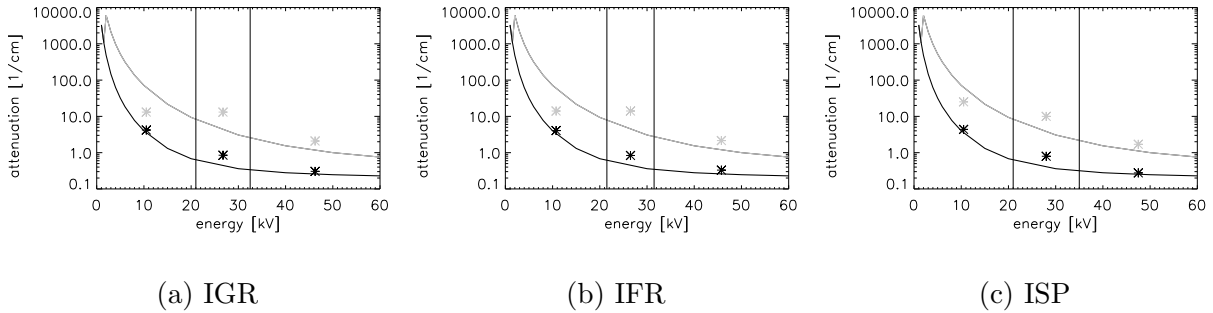


Figure 12. Comparison of the linear attenuation coefficients obtained by the three methods (stars) with the known attenuation curves (solid lines), where the grey and black lines represent the materials aluminum and PMMA, respectively. The obtained linear attenuation coefficients are plotted in the center of their corresponding energy bins (see text), which are delineated by the black vertical lines.

IFR, aim at reducing beam hardening artifacts in the reconstruction image by modeling the object function using relative density parameters. The ISP method, in which the relative  
 440 density parameter is omitted, is an iterative sinogram preprocessing method. For all three methods, the number of materials is considered to be known beforehand, but no information on the energy spectrum of the source-detector system or on the energy dependent attenuation coefficients of the materials is used. This is a significant practical advantage compared to most other beam hardening correction methods, which often require material and spec-  
 445 trum calibration. The three methods successfully reduce the cupping artifacts caused by

the beam polychromaticity in such a way that the reconstruction of each homogeneous region is to good accuracy homogeneous, even in case the segmentation of the preliminary reconstruction image is poor. In addition, the results show that the three methods are fairly robust for segmentation errors in the initial reconstruction, and allow visualizing small variations in the uniformity of the segments. For the simulated clinical scan in Fig. 6, the density variation was as large as a factor 2 in the bone.

A limitation of the method is that the reconstructed values cannot be interpreted as accurate estimates of the attenuation coefficients at a specific energy. Nevertheless, the results suggest that the relation between the estimated intensities of the energy bins and the reconstructed attenuations is compatible with the known material properties.

Our results also illustrated that, in spite of additional beam hardening effects introduced in the raw data, it can be favorable to perform the acquisition with little instead of strong hardware beam prefiltering so as to preserve the image contrast of low contrast objects.

As discussed in the introduction, another method that does not require prior knowledge on the spectrum and materials properties has been proposed by Krumm et al.<sup>26</sup>. That method is similar to the ISP method but fits the data with a generic polychromatic model, which depends on a much larger number of parameters than the ISP method since the latter uses the physical model in equation (17). The required number of parameters in the Krumm method grows exponentially with the number of materials, as opposed to linearly for the ISP algorithm. The price to pay is an increased computational complexity for the latter method.

The methods are general methods and their applicability is restricted to objects for which the spectral behavior within a single segment is sufficiently uniform. However, the results presented in this paper indicate that successful beam hardening artifact reduction can be obtained, even if there are relatively large density variations within a segment. Future work will be needed to investigate the influence of scatter, bow-tie filtering, systematic reconstruction errors, to compare the resolution and noise properties of the three proposed methods, and to determine the practical application field.

This work was financially supported by the Inter-University Attraction Poles Program 6-38 of the Belgian Science Policy, and by the SBO-project QUANTIVIAM (060819) of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). The authors would like to thank Arthur Van Gompel for the construction of the phantoms, and Elke Van de Castele and Evi Bongaers from SkyScan for providing the raw micro CT data.

### Appendix A: Update of the relative density in IGR

We present here the gradient descent algorithm used to calculate the relative density update step of the IGR method. The algorithm is derived by seeking a surrogate function  $\Phi_A(\mathbf{d}, \mathbf{d}^n)$  of the cost function in equation (9), which satisfies the usual conditions  $\Phi_A(\mathbf{d}, \mathbf{d}^n) \geq \Phi(\mathbf{d})$  and  $\Phi_A(\mathbf{d}^n, \mathbf{d}^n) = \Phi(\mathbf{d}^n)$ <sup>35,36</sup>. Note that all other parameters of the cost function are omitted here since they are kept constant during the relative density update step. The next iterate is then obtained as

$$\mathbf{d}^{n+1} = \arg \min_{\mathbf{d} \geq \mathbf{0}} \Phi_A(\mathbf{d}, \mathbf{d}^n) \quad (\text{A1})$$

and is easily calculated provided  $\Phi_A(\mathbf{d}, \mathbf{d}^n)$  has a simple structure. Due to the complexity of the cost function, we could only build a simple separable surrogate for the approximation of  $\Phi$  by a quadratic Taylor expansion:

$$\Phi(\mathbf{d}) \simeq \Phi(\mathbf{d}^n) + \sum_{j=1}^J \frac{\partial \Phi}{\partial d_j} \Big|_{\mathbf{d}^n} (d_j - d_j^n) + \frac{1}{2} \sum_{j=1}^J \sum_{h=1}^J \frac{\partial^2 \Phi}{\partial d_j \partial d_h} \Big|_{\mathbf{d}^n} (d_j - d_j^n)(d_h - d_h^n) \quad (\text{A2})$$

The first derivative of  $\Phi$  is given by:

$$\frac{\partial \Phi(\mathbf{d})}{\partial d_j} = \frac{2}{D} \sum_{e=1}^{E_m} \left[ \sum_m \mu_{m,e} s_{m,j} \right] \left[ \sum_i l_{i,j} \left( \ln \frac{I_{p,i}}{I_0} - \ln P_i \right) \cdot \frac{P_{i,e}}{P_i} \right] \quad (\text{A3})$$

with

$$P_{i,e} = I_e^F \exp \left( - \sum_m \mu_{m,e} \sum_j l_{i,j} d_j s_{m,j} \right) \quad (\text{A4})$$

$$P_i = \sum_e^{E_m} P_{i,e} \quad (\text{A5})$$

In addition we simplify the second derivative in (A2) by only keeping the dominant term of the second derivative, assuming that  $(\ln \frac{I_{p,i}}{I_0} - \ln P_i)$  can be neglected when we are close enough to the solution. With this approximation, the second derivative becomes:

$$\frac{\partial^2 \Phi(\mathbf{d})}{\partial d_j \partial d_h} \simeq H_{j,h}(\mathbf{d}) \doteq \frac{2}{D} \sum_e \left[ \sum_m \mu_{m,e} s_{m,j} \right] \cdot \sum_i l_{i,j} l_{i,h} P_{i,e} \frac{\sum_{e'} P_{i,e'} \sum_{m'} \mu_{m',e'} s_{m',h}}{P_i^2} \quad (\text{A6})$$

Because the RHS of (A6) is always positive and

$$2(d_j - d_j^n)(d_h - d_h^n) \leq (d_j - d_j^n)^2 + (d_h - d_h^n)^2 \quad (\text{A7})$$

485 it follows that:

$$\sum_j \sum_h H_{j,h}(\mathbf{d}^n) (d_j - d_j^n)(d_h - d_h^n) \leq \sum_j \sum_h H_{j,h}(\mathbf{d}^n) (d_j - d_j^n)^2 \quad (\text{A8})$$

This allows to introduce the surrogate  $\Phi_A$

$$\Phi_A(\mathbf{d}, \mathbf{d}^n) = \Phi(\mathbf{d}^n) + \sum_j \left. \frac{\partial \Phi}{\partial d_j} \right|_{\mathbf{d}^n} (d_j - d_j^n) + \frac{1}{2} \sum_j \sum_h H_{j,h}(\mathbf{d}^n) (d_j - d_j^n)^2. \quad (\text{A9})$$

This function satisfies  $\Phi_A(\mathbf{d}^n, \mathbf{d}^n) = \Phi(\mathbf{d}^n)$  and (at the approximations described above),  $\Phi_A(\mathbf{d}^n, \mathbf{d}^n) \geq \Phi(\mathbf{d}^n)$ . Minimizing  $\Phi_A(\mathbf{d}, \mathbf{d}^n)$  is trivial since the variables are separated. Setting the derivative of  $\Phi_A$  with respect to  $d_j$  to zero yields

$$d_j = d_j^n - \frac{\left. \frac{\partial \Phi}{\partial d_j} \right|_{\mathbf{d}^n}}{\sum_h H_{j,h}(\mathbf{d}^n)}, \quad (\text{A10})$$

which corresponds to a gradient descent with a diagonal pre-conditionner  $1/\sum_h H_{j,h}(\mathbf{d}^n)$ . The second derivative in (Eq. (A6)) contains one backprojection for each energy bin. To limit the number of backprojections we multiply each term in the second derivative by  $\frac{P_i}{P_{i,e}} > 1$ , which results in only one backprojection. This approximation does not lead to instability  
490 because it only decreases the step size. Adding a final non-negativity constraint, the ISF update step becomes:

$$d_j^{n+1} = \left[ d_j^n - \frac{\sum_e \left[ \sum_m \mu_{m,e} s_{m,j} \right] \left[ \sum_i l_{i,j} \left( \ln \frac{I_{p,i}}{I_0} - \ln P_i \right) \cdot \frac{P_{i,e}}{P_i} \right]}{\sum_e \left[ \sum_m \mu_{m,e} s_{m,j} \right] \cdot \sum_i l_{i,j} \frac{1}{P_i} \sum_{e'} P_{i,e'} \sum_{m'} \mu_{m',e'} \left( \sum_h l_{i,h} s_{m',h} \right)} \right]_+ \quad (\text{A11})$$

with  $[x]_+ = x$  if  $x \geq 0$  and  $[x]_+ = 0$  if  $x < 0$ .

## REFERENCES

- <sup>1</sup>J. Stonestrom, R. E. Alvarez, and A. Mackowski, “A framework for spectral artifact  
495 corrections in X-ray CT,” *IEEE Transactions on Biomedical Engineering*, **28**, 128–141  
(1981).
- <sup>2</sup>R. Alvarez and A. Macowski, “Energy-selective Reconstructions in X-ray Computerized  
Tomography,” *Physics in Medicine and Biology*, **21**, 733–744 (1976).
- <sup>3</sup>F. Kelcz, P. Joseph, and S. Hilal, “Noise considerations in dual energy CT scanning,”  
500 *Medical Physics*, **6**, 418–425 (1979).
- <sup>4</sup>W. A. Kalender, D. Felsenberg, H. K. Genant, M. Fischer, J. Dequeker, and F. Reeve,  
“The European Spine Phantom a tool for standardization and quality control in spinal  
bone mineral measurements by DXA and QCT,” *European Journal of Radiology*, **20**,  
83–92 (1995).
- <sup>5</sup>G. Lu, S. Wu, B. Yeh, and L. Zhang, “Dual-energy computed tomography in pulmonary  
505 embolism,” *British Journal of Radiology*, ID 16337436 (2010).
- <sup>6</sup>G. Bastarrika, Y. Lee, W. Huda, B. Ruzsics, P. Costello, and U. Schoepf, “CT of Coronary  
Artery Disease,” *Radiology*, **253**, 317–338 (2009).
- <sup>7</sup>I. Elbakri and J. Fessler, “Segmentation-free statistical image reconstruction for polyener-  
510 getic x-ray computed tomography with experimental validation,” *Physics in Medicine and  
Biology*, **48**, 2453–2477 (2003).
- <sup>8</sup>N. Menvielle, Y. Goussard, D. Orban, and G. Soulez, “Reduction of beam-hardening  
artifacts in x-ray ct,” *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS  
2005. 27th Annual International Conference of the, 1865–1868* (2005).
- <sup>9</sup>B. De Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, “An Iterative Maximum-  
515 Likelihood Polychromatic Algorithm for CT,” *IEEE Trans. Med. Imaging*, **20**, 999–1008  
(2001).
- <sup>10</sup>J. O’Sullivan and J. Benac, “Alternating Minimization Algorithms for Transmission To-  
mography,” *IEEE Transactions on Medical Imaging*, **26**, 283–297 (2007).
- <sup>11</sup>E. Van de Castele, D. Van Dyck, J. Sijbers, and E. Raman, “An energy-based beam  
520 hardening model in tomography,” *Physics in Medicine and Biology*, **47**, 4181–4190 (2002).
- <sup>12</sup>G. Herman, “Correction for beam hardening in computed tomography,” *Physics in  
Medicine and Biology*, **24**, 81–106 (1979).

- <sup>13</sup>M. Kachelriess, K. Sourbelle, and W. Kalender, “Empirical cupping correction: a first-  
525 order raw data precorrection for cone-beam computed tomography,” *Medical Physics*, **33**,  
1269–1274 (2006).
- <sup>14</sup>Y. Censor, T. Elfving, and G. Herman, “A method of iterative data refinement and its  
applications,” *Mathematical Methods in Applied Science*, **7**, 108–123 (1985).
- <sup>15</sup>P. Joseph and R. Spital, “A method for correcting bone induced artifacts in computed  
530 tomography scanners,” *Journal of Computer Assisted Tomography*, **2**, 100–109 (1978).
- <sup>16</sup>G. Herman and S. Trivedi, “A comparative study of two postreconstruction beam hard-  
ening correction methods,” *IEEE Transactions on Medical Imaging*, **2**, 128–135 (1983).
- <sup>17</sup>P. Kijewski and B. Bjärngard, “Correction for beam hardening in computed tomography,”  
*Medical Physics*, **5**, 209–214 (1978).
- <sup>18</sup>P. Joseph and C. Ruth, “A method for simultaneous correction of spectrum hardening  
535 artifacts in CT images containing both bone and iodine,” *Medical Physics*, **24**, 1629–1634  
(1997).
- <sup>19</sup>J. Hsieh, R. Molthen, C. Dawson, and R. Johnson, “An iterative approach to the beam  
hardening correction in cone beam CT,” *Medical Physics*, **27**, 23–29 (2000).
- <sup>20</sup>V. Vedula and P. Munshi, “An improved algorithm for beam-hardening corrections in  
540 experimental X-ray tomography,” *NDT&E International*, **41**, 25–31 (2008).
- <sup>21</sup>J. Meagher, C. Mote, and H. Skinner, “CT image correction for beam hardening using  
simulated projection data,” *IEEE Transactions on Nuclear Science*, **37**, 1520–1524 (1990).
- <sup>22</sup>T. Fuchs, *Strahlauhärtungskorrekturen in der computer-tomographie*, Ph.D. thesis,  
545 Friedrich-Alexander Universität, Erlangen-Nürnberg, Germany (1998).
- <sup>23</sup>E. Van de Casteele, D. Van Dyck, J. Sijbers, and E. Raman, “A model-based correction  
method for beam hardening artefacts in X-ray microtomography,” *Journal of X-ray Science  
and Technology*, **12**, 43–57 (2004).
- <sup>24</sup>H. Gao, L. Zhang, Z. Chen, Y. Xing, and L. Shuanglei, “Beam hardening correction  
550 for middle-energy industrial computerized tomography,” *IEEE Transactions on Nuclear  
Science*, **53**, 2796–2807 (2006).
- <sup>25</sup>X. Mou, S. Tang, and H. Yu, “Comparison on Beam Hardening Correction of CT Based  
on H-L Consistency and Normal Water Phantom Experiment,” in *SPIE Medical Imaging:  
Image Processing*, Vol. 6318 (2006) p. 63181V.

- 555 <sup>26</sup>M. Krumm, S. Kasperl, and M. Franz, “Reducing non-linear artifacts of multi-material objects in industrial 3D computed tomography,” *Nondestructive Testing and Evaluation*, **41**, 242–251 (2008).
- <sup>27</sup>O. Nalcioglu and R. Lou, “Post-reconstruction method for beam hardening in computerised tomography,” *Physics in Medicine and Biology*, **24**, 333–340 (1979).
- 560 <sup>28</sup>L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *Journal of Optical Society of America*, **1**, 612–619 (1984).
- <sup>29</sup>A. Katsevitch, “Theoretically exact filtered backprojection-type inversion algorithm for spiral CT,” *SIAM Journal of Applied Mathematics*, **62**, 2012–2026 (2002).
- <sup>30</sup>J. A. Hartigan, *Clustering Algorithms* (Wiley, 1975).
- 565 <sup>31</sup>H. Hudson and R. Larkin, “Accelerated image reconstruction using ordered subsets of projection data,” *IEEE Transactions on Medical Imaging*, **13**, 601–609 (1994).
- <sup>32</sup>R. Birch, M. Marshall, and G. Adran, “Catalogue of Spectral Data for Diagnostic X-Rays,” The Hospital Physicists’ Association, London (1979).
- <sup>33</sup>ICRU Report No. 44, “Tissue substitutes in radiation dosimetry and measurement,” International Commission on Radiation Units and Measurements, Bethesda, MD, USA (1989).
- 570 <sup>34</sup>K. Batenburg and J. Sijbers, “Optimal Threshold Selection for Tomogram Segmentation by Projection Distance Minimization,” *IEEE Transactions on Medical Imaging*, **28**, 676–686 (2009).
- <sup>35</sup>A. De Pierro, “A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography,” *IEEE Transactions on Medical Imaging*, **14**, 132–137
- 575 (1995).
- <sup>36</sup>J. Nuyts, S. Stroobants, P. Dupont, S. Vleugels, P. Flamen, and L. Mortelmans, “Reducing Loss of Image Quality Because of the Attenuation Artifact in Uncorrected PET Whole-Body Images,” *Journal of Nuclear Medicine*, **43**, 1054–1062 (2002).