Published in final edited form as:

Phys Med Biol. 2016 May 07; 61(9): 3387–3406. doi:10.1088/0031-9155/61/9/3387.

Artifact reduction in short-scan CBCT by use of optimization-based reconstruction

Zheng Zhang¹, Xiao Han¹, Erik Pearson^{2,3}, Charles Pelizzari², Emil Y Sidky¹, and Xiaochuan Pan^{1,2}

¹Department of Radiology, The University of Chicago, Chicago, IL, USA

²Department of Radiation and Cellular Oncology, The University of Chicago, Chicago, IL, USA

Abstract

Increasing interest in optimization-based reconstruction in research on, and applications of, cone-beam computed tomography (CBCT) exists because it has been shown to have the potential to reduce artifacts observed in reconstructions obtained with the Feldkamp–Davis–Kress (FDK) algorithm (or its variants), which is used extensively for image reconstruction in current CBCT applications. In this work, we carried out a study on optimization-based reconstruction for possible reduction of artifacts in FDK reconstruction specifically from short-scan CBCT data. The investigation includes a set of optimization programs such as the image-total-variation (TV)-constrained data-divergency minimization, data-weighting matrices such as the Parker weighting matrix, and objects of practical interest for demonstrating and assessing the degree of artifact reduction. Results of investigative work reveal that appropriately designed optimization-based reconstruction, including the image-TV-constrained reconstruction, can reduce significant artifacts observed in FDK reconstruction in CBCT with a short-scan configuration.

Keywords

optimization; image reconstruction; cone-beam CT; iterative algorithm

1. Introduction

Cone-beam computed tomography (CBCT) has been widely used for guidance, monitoring, and assessment tasks in interventional (Wiesent *et al* 2000, Lauritsch *et al* 2006, Orth *et al* 2008, Wallace *et al* 2008), surgical (Grass *et al* 1999, Hott *et al* 2004, Siewerdsen *et al* 2005, De Vos *et al* 2009), and therapeutic applications (Jaffray *et al* 2002, Oldham *et al* 2005, Smitsmans *et al* 2005, Grills *et al* 2008). Due to workflow and radiation-dose considerations, it is often desirable to scan a patient with the shortest possible rotation scan that provides sufficient data for reconstruction. With the commonly used Feldkamp–Davis–Kress (FDK) algorithm (or its variants) (Feldkamp *et al* 1984, Johnson *et al* 1998), this is 180°+ the-full-fan-angle, which is typically referred to as the short-scan angular range (Parker 1982). Although the FDK algorithm can provide images of high practical utility for a wide class of

³Current Address: Techna Institute and Princess Margaret Cancer Centre of the University Health Network, Toronto, ON M5R 1S8, Canada

applications, it has also been recognized, however, to yield reconstructions with noticeable artifacts especially in the off-middle planes for short-scan CBCT with data insufficiency and inconsistency (Dennerlein *et al* 2008, Maaß *et al* 2010).

Optimization-based reconstruction (Delaney *et al* 1998, Elbakri and Fessler 2002, Sidky and Pan 2008, Tang *et al* 2009, Pan *et al* 2009, Bian *et al* 2010, Stsepankou *et al* 2012, Nien and Fessler 2014, Wang *et al* 2014) remains the focus of current research on CBCT algorithms. Evidence seems to suggest that appropriately designed optimization-based algorithms may be less susceptible to data inconsistency and insufficiency than analytic algorithms such as the FDK algorithm, thus possibly yielding reconstructions with reduced artifacts (Parker 1982, Dennerlein *et al* 2008). In this work, we carry out a study on optimization-based reconstructions from short-scan CBCT data with an emphasis on demonstrating whether different designs of optimization programs of potential practical interest impact the reduction of artifacts observed in the FDK reconstruction.

Based upon the CBCT-imaging model for analytic algorithms, it can be shown that the data function in a short-scan CBCT possesses ‘redundancy’ within the middle plane, i.e. the plane containing the source trajectory. A weighting function such as the Parker weighting function (Parker 1982) is used widely for normalizing the data-function redundancy so that the FDK algorithm can, when applied to the weighted data-function, yield accurate images within the middle plane. It is well-known that different weighting functions may lead to FDK reconstructions with different degrees of artifacts especially away from the middle plane (Grass *et al* 2001, Tang *et al* 2005, Mori *et al* 2006). Therefore, we also investigate possible impact of different weighting functions, especially those that normalize the data-function redundancy, on optimization-based reconstruction in short-scan CBCT.

This paper starts with an introduction in section 1, followed by the description of optimization reconstruction in section 2, and of study design in section 3. Results of real-data studies and final remarks are presented in sections 4 and 5, respectively.

2. Methods

The imaging model used in the optimization-based reconstruction considered is given by

$$\mathbf{g} = \mathcal{H} \mathbf{f}, \quad (1)$$

where vector $\mathbf{g}$ of size M denotes discrete model-data, vector $\mathbf{f}$ of size N the discrete image of N voxels, and $\mathcal{H}$ the system matrix of size $M \times N$. In this work, an entry in $\mathbf{g}$ or $\mathbf{f}$ represents the data or image value within a detector bin or image voxel; and an element in system matrix $\mathcal{H}$ represents an intersection of an x-ray detected with a voxel in the image array. The imaging model in equation (1), linking a discrete-image vector to a discrete-data vector, is referred to as a discrete-to-discrete (D–D) model. We denote experimentally measured data in practical CBCT imaging as vector $\mathbf{g}_m$ of size M in which an entry represents a particular data sample collected. In contrast, the imaging model upon which analytic algorithms such as the FDK algorithm are developed is referred to a continuous-to-

continuous (C-C) model because it maps an object function of continuous variables to a data function of continuous variables.

2.1. Design consideration of optimization programs

In optimization-based reconstruction, an image is obtained by solving the optimization program that specifies the image. In this work, we consider optimization programs of form

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} \Phi(\mathcal{W} \mathbf{g}_m, \mathcal{W} \mathcal{H} \mathbf{f}) \quad \text{s. t.} \quad \Psi(\mathbf{f}) \leq \alpha \quad \text{and} \quad f_j \geq 0, \quad (2)$$

where $\Phi(\mathcal{W} \mathbf{g}_m, \mathcal{W} \mathcal{H} \mathbf{f})$, referred to data fidelity in the work, denotes a measure of inconsistency between measured data $\mathbf{g}_m$ and model data $\mathbf{g}$, weighted often by a matrix $\mathcal{W}$ of size $M \times M$, $\Psi(\mathbf{f})$ an image constraint function, α constraint parameter, and f_j the j th entry of vector $\mathbf{f}$, $j = 1, 2, \dots, N$.

2.1.1. Data fidelities—Several data fidelities and image constraints of potential practical interest are considered in the work. Specifically, the data fidelities are given by

a. *data- ℓ_1 fidelity*

$$D_{\ell_1}(\mathbf{f}) = \|\mathcal{W} \mathbf{g}_m - \mathcal{W} \mathcal{H} \mathbf{f}\|_1, \quad (3)$$

where $\|\cdot\|_1$ denotes ℓ_1 -norm;

b. *data- ℓ_2 fidelity*

$$D_{\ell_2}(\mathbf{f}) = \|\mathcal{W} \mathbf{g}_m - \mathcal{W} \mathcal{H} \mathbf{f}\|_2^2, \quad (4)$$

where $\|\cdot\|_2^2$ denotes the square of ℓ_2 -norm; and

c. *data-Kullback-Leibler (KL) fidelity*

$$D_{\text{KL}}(\mathbf{f}) = \sum_{i=1}^M [-\mathcal{W} \mathcal{H} \mathbf{f} + \mathcal{W} \mathbf{g}_m + \mathcal{W} \mathbf{g}_m \ln(\mathcal{W} \mathbf{g}_m) - \mathcal{W} \mathbf{g}_m \ln(\mathcal{W} \mathcal{H} \mathbf{f})]_i \quad (5)$$

where i is the index of the i th entry of data vector $\mathbf{g}_m$, and the entries in vector $(\mathcal{W} \mathcal{H} \mathbf{f})$ that are smaller than 10^{-20} are replaced with 10^{-20} . In contrast to practical FDK reconstruction in which measured data $\mathbf{g}_m$ is pre-conditioned by a diagonal weighting matrix of size $M \times M$, a weighting matrix in an optimization-based reconstruction is multiplied with both measured data $\mathbf{g}_m$ and model data $\mathbf{g}$, thus modulating inconsistency between $\mathbf{g}_m$ and $\mathbf{g}$, instead of $\mathbf{g}_m$ alone.

2.1.2. Image constraints—The image constraints considered are given by

a. *image- ℓ_1 constraint*

$$\|\mathbf{f}\|_{\ell_1} = \sum_{j=1}^N |f_j|; \quad (6)$$

b. *image- ℓ_2 constraint*

$$\|\mathbf{f}\|_{\ell_2} = \sqrt{\sum_{j=1}^N f_j^2}; \quad (7)$$

and

c. *image-TV constraint*

$$\|\mathbf{f}\|_{TV} = \|\nabla \mathbf{f}\|_{MAG}, \quad (8)$$

where $\|\nabla \mathbf{f}\|_{MAG}$ denotes the magnitude image of the spatial gradient of $\mathbf{f}$ defined above Eq. (A.7) in the Appendix.

2.2. Optimization program with image-TV constraint

Using the data fidelities and image constraints defined above to replace, respectively, Φ and Ψ in equation (2), we form and study a number of optimization programs in the work, and, for discussion convenience, dubbed with respective symbols summarized in table 1. Image-constraint parameter α corresponding to image constraints $\|\mathbf{f}\|_{\ell_1}$, $\|\mathbf{f}\|_{\ell_2}$, or $\|\mathbf{f}\|_{TV}$ is denoted, respectively, as ℓ_1 , ℓ_2 , or t_1 , as shown in the last column of table 1. For the purpose of clarity and brevity, without loss of generality, we give below explicit forms of optimization programs with the image-TV constraint, shown in row 5 of table 1. In section 4.3 we also present results of two additional optimization programs using image- ℓ_1 and image- ℓ_2 constraints, respectively.

2.2.1. Program $D_{\ell_1}-\|\mathbf{f}\|_{TV}$: image-TV-constrained data- ℓ_1 minimization—

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} D_{\ell_1}(\mathbf{f}) \quad \text{s. t.} \quad \|\mathbf{f}\|_{TV} \leq t_1 \quad \text{and} \quad f_j \geq 0, \quad (9)$$

where $t_1 \geq 0$ denotes a pre-selected image-constraint parameter.

2.2.2. Program $D_{\ell_2}-\|\mathbf{f}\|_{TV}$: image-TV-constrained data- ℓ_2 minimization

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} D_{\ell_2}(\mathbf{f}) \quad \text{s. t.} \quad \|\mathbf{f}\|_{TV} \leq t_1 \quad \text{and} \quad f_j \geq 0. \quad (10)$$

2.2.3. Program $D_{\text{KL}}\|\mathbf{f}\|_{\text{TV}}$: image-TV-constrained data-KL minimization

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} D_{\text{KL}}(\mathbf{f}) \quad \text{s. t.} \quad \|\mathbf{f}\|_{\text{TV}} \leq t_1 \quad \text{and} \quad f_j \geq 0. \quad (11)$$

2.2.4. Program parameters—Even if the mathematical form of an optimization program is given, parameters are needed also for completely specifying the program. As mentioned above, system matrix $\mathcal{H}$ needs to be determined, and, in this work, each of its elements is selected as the intersection length of an x-ray with an image voxel. The selection of weighting matrix $\mathcal{W}$, which modulates data inconsistency, is discussed below in section 2.3. Additionally, the programs in equations (9)–(11) also depend upon voxel size of the image array and image-TV parameter t_1 . When image- ℓ_1 or image- ℓ_2 , instead of image-TV, is used as an image constraint, the optimization program depends upon constraint parameter I_1 or I_2 (see row 3 and 4 in table 1). Different selections of parameters specify generally different solutions (i.e. different reconstructions). We discuss in section 3.2 below the determination of the program parameter in real-data studies.

2.3. Weighting matrix

In short-scan CBCT, let u_m and v_m denote detector sizes in directions tangential to the circular trajectory and perpendicular to the middle plane; and u and v are continuous variables, depicting the Cartesian coordinates of a point on the detector plane, satisfying $|u| \leq u_m$ and $|v| \leq v_m$. At continuous projection angle ϕ , the weighting function normalizing data-function redundancy within the middle plane satisfies

$$w(u, v, \phi) + w(-u, v, \phi + \pi + 2\gamma) = 1, \quad (12)$$

where $\gamma = \operatorname{atan}(\frac{v}{L})$, and L the distance from the source to the detector surface. Even though the weighting function is designed for the middle plane, it is used also for normalizing approximate data-function redundancy within the off-middle planes (i.e. $v \neq 0$) in circular CBCT. An additional condition is considered often on the weighting-function smoothness (Parker 1982) for FDK reconstruction:

$$\begin{cases} \left. \frac{\partial w(u, v, \phi)}{\partial \phi} \right|_{\phi=2\gamma_m-2\gamma} = \left. \frac{\partial w(u, v, \phi)}{\partial \phi} \right|_{\phi=\pi-2\gamma} = 0 \\ \left. \frac{\partial w(u, v, \phi)}{\partial u} \right|_{\phi=2\gamma_m-2\gamma} = \left. \frac{\partial w(u, v, \phi)}{\partial u} \right|_{\phi=\pi-2\gamma} = 0, \end{cases} \quad (13)$$

where $\gamma_m = \operatorname{atan}(\frac{u_m}{L})$. It has been shown that the smoothness condition is critical for minimizing sampling artifacts in FDK reconstruction from discrete data. The Parker weighting function (Parker 1982)

$$\begin{cases} w(u, v, \phi) = \sin^2\left(\frac{\pi}{4} \frac{\phi}{\gamma_m - \gamma}\right) & 0 \leq \phi < 2\gamma_m - 2\gamma \\ w(u, v, \phi) = 1 & 2\gamma_m - 2\gamma \leq \phi < \pi - 2\gamma \\ w(u, v, \phi) = \sin^2\left(\frac{\pi}{4} \frac{\pi + 2\gamma_m - \phi}{\gamma_m + \gamma}\right) & \pi - 2\gamma \leq \phi < \pi + 2\gamma_m \\ w(u, v, \phi) = 0 & \pi + 2\gamma_m \leq \phi < 2\pi. \end{cases} \quad (14)$$

satisfies conditions in equations (12) and (13), and is used widely for image reconstruction in short-scan CBCT.

In analytic reconstruction, a weighting function is used for normalizing the data-function redundancy within the middle plane of short-scan CBCT. In real-data imaging, only discrete data $\mathbf{g}_m$ of size M can be measured, and, a diagonal *weighting matrix* is constructed thus by using M values, in a *concatenated form*, calculated from a weighting function satisfying equations (12) and (13) at M discrete points in data space (u, v, ϕ) at which $\mathbf{g}_m$ are sampled. The FDK algorithm then reconstructs an image from the measured, discrete data weighted by the diagonal weighting matrix. On the other hand, optimization-based reconstruction can accommodate weighting matrices that are designed based upon physical, statistical, and/or mathematical motivations other than data-redundancy normalization. In an attempt to compare with FDK reconstructions that use the diagonal weighting matrix obtained from the Parker weighting function, the same weighting matrix is used in optimization programs considered in the work, while additional weighting matrices that are designed without consideration of normalizing the data-function redundancy are also studied in section 4.3 below.

2.4. Algorithm derivation and validation

2.4.1. Algorithm derivation—The optimization programs considered in table 1 are convex, and they can thus be solved by use of the primal-dual algorithms. In the work, we use the primal-dual algorithm developed by Chambolle and Pock (CP) (Chambolle and Pock 2011, Sidky *et al* 2012) because it appears to be capable of handling effectively system matrix $\mathcal{H}$ of large size of practical CT interest (Sidky *et al* 2014), and because it solves not only smooth, but also non-smooth, programs listed in table 1. The use of a single algorithm for solving all of the optimization programs may also avoid any reconstruction differences as a result of using different algorithms for solving different optimization programs. For the programs in table 1, we have derived explicitly the proximal-mapping steps in the CP algorithm, i.e. the update procedures in lines 9 and 10 shown in the algorithm pseudo-code in the appendix, thus allowing accurate and efficient algorithm implementation.

2.4.2. Algorithm validation—In an attempt to verify first the implementation of the CP algorithm, and its appropriate applications to solving the optimization programs, we have performed, for each of the optimization programs considered, an inverse-crime study in which model data were generated by use of a system matrix from a known truth image, and the same system matrix is used also in reconstruction. Therefore, in each inverse-crime study, program parameters such as image-voxel size and image-constraint parameters (i.e. t_1 , l_1 , or l_2) are known (Bian *et al* 2014, Han *et al* 2015) with which the CP algorithm was

validated quantitatively to converge to the truth image numerically through solving the optimization program. The results of the necessary validation study are not shown, because the focus of the work is on investigating possible reduction of artifacts observed in FDK reconstruction by use of optimization-based reconstruction.

3. Study design and materials

3.1. Data collection and FDK reference

3.1.1. Short-scan CBCT data—We collected data from two physical phantoms, the RANDO man and the Catphan 504 (The Phantom Laboratory, Salem NY), using a Varian on-board-imager (OBI) on a Trilogy LINAC (Varian Medical System, Palo Alto CA). The two phantoms are selected because they contain both anthropomorphic structures in RANDO and geometric, quantitative structures in the Catphan that can be used for evaluating the impact of reconstruction algorithms. The OBI CBCT system consists of a keV x-ray source and a flat-panel detector of sizes $2u_m = 39.73$ cm and $2v_m = 29.80$ cm, composing of 1024×768 identical square-shaped-detector bins of size 0.0388 cm. The distances from the source to the rotation axis, and to the flat-detector surface, are $S = 100$ cm and $L = 150$ cm, respectively. Data were acquired at 353 and 347 views, respectively, for the Rando and Catphan phantoms, uniformly distributed over a short-scan angular range of 196 degrees. Therefore, the sizes of data vectors considered are $M = 1024 \times 768 \times 353$ and $1024 \times 768 \times 347$ for the cases involving, respectively, the Rando and Catphan phantoms.

3.1.2. FDK reference—For comparison, we also carried out reconstructions by using the FDK algorithm, with a Hanning filter and a cutoff at 0.5 of the Nyquist frequency, from weighted real-data $\mathcal{W}g_m$. For revealing details of artifact reduction, soft-tissue display windows are used for reconstructions; and, as shown in figure 1, two regions of interest (ROIs) enclosed by rectangles within each of the FDK reconstructions of the Rando and Catphan phantoms are selected for showing ROI images in zoomed-in views.

3.2. Program-parameter selection

3.2.1. Weighting matrix and image-voxel size—Results below were obtained with a diagonal weighting matrix that is formed from discrete values of the Parker weighting function, except for the two cases in section 4.3, in which different weighting matrices were used. In a real-data study, image-voxel size and image-constraint parameter need to be estimated, because only measured data g_m are available experimentally (Bian *et al* 2014, Han *et al* 2015). We use in the Rando and Catphan studies image arrays of sizes $N = 512 \times 512 \times 451$ and $512 \times 512 \times 239$, respectively, consisting of cubic voxels of size identical to that used in clinical IGRT CBCT imaging, which is 1.258 times of the detector-bin size of OBI CBCT scanner used for collecting data.

3.2.2. Image-constraint parameter—We performed multiple reconstructions from each data set collected, with different t_1 values. Based upon a visual inspection of these reconstructions, we then select t_1 that yields a reconstruction with reduced artifacts comparing to the FDK image, yet without worsened, or with improved, contrast and spatial resolution. Such a way of selecting t_1 is admittedly subjective, but it seems sufficient for the

purpose of the work which is qualitatively demonstrating the potential of image-TV-constrained reconstruction for reduction of artifacts observed in FDK reconstruction from short-scan CBCT data. Similarly, we determined image-constraint parameters l_1 or l_2 in optimization programs using image- l_1 or image- l_2 constraint, by performing reconstructions with varying values of l_1 or of l_2 instead, and then selecting l_1 or l_2 that yields reconstruction with visually least artifacts. The specific selection of an image-constraint parameter is generally data- and object-dependent.

3.2.3. Practical convergence conditions—The CP algorithm has been shown to converge mathematically to solutions specified by a convex optimization program, including the optimization programs considered, but with an infinite number of iterations, which cannot be achieved in practice. Instead, using the mathematical convergence conditions as a guide, we developed practical convergence conditions such that they can be achieved within a finite number of iterations (Bian *et al* 2014, Han *et al* 2015). Of course, practical convergence conditions should be interpreted as additional program parameters because they play now, along with other program parameters, a role in specifying feasible solutions. We consider in real-data studies three practical convergent conditions to be satisfied collectively as functions of iteration numbers: (1) the image- l_1 , image- l_2 , or image-TV constraint of the reconstructed image arrives at its corresponding, designed image constraint value; (2) the data fidelity data- l_2 , data- l_1 , or data-KL sufficiently plateau; and (3) the conditional primal-dual gap (cPD) (Sidky *et al* 2012) continuously decreases when conditions (1) and (2) are reached. (The mathematical cPD condition on the CP-algorithm convergence is that the cPD approaches zero. Depending upon the property of the system matrix, data condition, optimization program, and computer precision, it is often unrealistic to achieve numerically the cPD condition in a practical reconstruction. As the results below show, however, the three practical convergence conditions defined above are able to yield reconstructions of practical interest.)

4. Study results

From the perspective of a C–C data model for circular CBCT, short-scan data sufficient for accurate reconstruction within the middle plane are insufficient for non-middle planes. In a real-data study, data inconsistency, including noise and scatter, can lead to artifacts within the middle plane. However, for non-middle planes, artifacts result from not only data inconsistency but also data insufficiency. The results below highlight three major studies: (1) We first study optimization programs in table 1 for reconstructions within the middle plane, which contains no data-insufficiency artifacts. (2) Based upon the study results, we select the optimization programs, which are shown to perform well based upon visual inspection, for further investigation of their potential in reduction of artifacts observed within non-middle planes of FDK reconstructions caused by data insufficiency. (3) Additional studies are performed on how the artifact reduction is affected by the selection of additional optimization programs, and weighting matrices.

4.1. Impact of optimization programs on reconstructions

4.1.1. Reconstructions based upon optimization programs in row 2 of table 1

—We reconstructed images from Rando data by solving the optimization programs in row 2 of table 1, and show the convergent reconstructions in row 1 of figure 2. As expected, because the optimization programs contain only data fidelities (except for an image positivity constraint,) the convergent reconstructions are all highly noisy, with image textures considerably different than that in the FDK reconstruction. Visual differences exist among the reconstructions because different designs of optimization programs necessarily respond differently to data inconsistencies such as data noise. It can be observed that reconstruction of program D_{l_1} with data- l_1 fidelity appears with sprinkled-salt texture, and is noisier than those obtained with programs D_{l_2} and D_{KL} containing data- l_2 and data-KL fidelities. Results similar to Rando reconstructions have been obtained also for Catphan reconstructions, which are not shown. The study suggests the need of adequate image constraints for image reconstruction from real data.

4.1.2. Reconstructions based upon optimization programs in rows 3–5 of table 1

—We then performed reconstructions from Rando data by solving the optimization programs in rows 3–5 of table 1, each of which containing a data fidelity and an image constraint (in addition to the image positivity constraint.) As discussed in section 3.2, for each program studied, we conducted a number of reconstructions with different values of its image-constraint parameter; and, based upon visual inspection of the reconstructions, we selected a final value and used it to yield convergent reconstructions, as shown in rows 2–4 of figure 2, according to the practical convergent conditions discussed in section 3.2. Result in row 2 reveals that reconstructions with image- l_1 constraint remain of noisy, sprinkled-salt texture. This is not surprising because image- l_1 constraint encourages sparse solutions, while the Rando-phantom image is substantially non-sparse. Conversely, as image- l_2 constraint promotes non-sparse solutions, optimization programs with image- l_2 constraint thus lead to reconstructions of the Rando phantom with visually reasonable appearance, as shown in row 3 of figure 2. On the other hand, reconstructions with an image-TV constraint in row 4 of figure 2 appear comparable to those obtained with image- l_2 , but with bony structures of enhanced spatial resolution. For cases considered, image constraints appear to impact reconstruction textures more significantly than do data fidelities, as reconstructions vary visually more significantly within each column than within each row. Observations similar to those for reconstructions of the Rando phantom can also be made for reconstructions of the Catphan phantom.

4.2. Optimization-based reconstructions and artifact reduction

We investigate below optimization-based reconstructions for possible reduction of artifacts within non-middle planes in FDK reconstructions.

4.2.1. Reconstructions—As figure 2 shows, optimization programs with image-TV constraint appear to yield reconstructions with visually reasonable spatial and contrast resolution as well as image texture. Therefore, for brevity, without loss of generality, we present below results of reconstructions specified by optimization programs with image-TV constraint in row 5 of table 1 (i.e. equations (9)–(11)). In the study, using the strategy

described in section 3.2, we determined image parameter t_1 for the Rando data, and display in figure 3 convergent reconstructions, accompanied by the corresponding FDK reconstructions. For revealing reconstruction details, zoomed-view images within the ROIs defined in figure 1 are depicted immediately below each corresponding reconstruction. Also, in figure 4, we display convergent reconstructions of the Catphan phantom obtained by solving the programs with image-TV constraint in row 5 of table 1. Similar results have also been obtained for reconstructions within other transverse, coronal, and sagittal slices of the two phantoms, which are not shown for brevity.

4.2.2. Artifact reduction—Results in figures 3 and 4 indicate that reconstructions specified by the programs with image-TV constraint in row 5 of table 1 appear to suppress effectively artifacts observed in FDK reconstructions. In particular, as a result of artifact reduction, low-contrast structures, highlighted by arrows, within the Rando phantom can now be observed in figure 3. Some residual artifacts remain visible in the reconstructions due to data insufficiency and inconsistency, such as beam-hardening and scatter, which are not included in the imaging model. It is also interesting to notice that image texture obtained with the programs appear differently, because they respond differently to data inconsistency and insufficiency. Reconstructions with image-TV constraint seem to be robust relative to different data fidelities in terms of compensating for the artifacts. Again, remarks similar to those for the Rando reconstructions above can be made for the Catphan reconstructions. Furthermore, inspection of figures 3 and 4 indicates that the performance of the optimization programs considered appears to be comparably effective for objects such as the Rando and Catphan phantoms with distinct anatomies.

4.2.3. Reconstruction evolution as a function of iterations—The discussion above focuses on convergent reconstructions satisfying the practical convergent conditions, obtained at about 1090 iterations for the Rando phantom, and about 950 iterations for the Catphan phantom, respectively. For demonstrating the reconstruction convergence, we show in figure 5 convergence metrics, including data- ℓ_2 , image-TV, and cPD, as functions of iterations for the Rando reconstruction in solving program $D_{\text{KL}}-\|\mathbf{f}\|_{\text{TV}}$ in equation (10). It can be observed that the convergence metrics plateau considerably before 1090 iterations. It is possible, however, that artifacts are compensated effectively for in reconstructions at earlier iterations. In figure 6, we display reconstructions of the Rando phantom at iterations 30, 50, and 150, and of the Catphan phantom at iterations 10, 30, and 100, as well as the corresponding convergent reconstructions, i.e. the final reconstructions, shown in column 4. It can be observed that Rando and Catphan reconstructions at, e.g. about iterations 150 already resemble visually their respective convergent reconstructions in terms of effective reduction of artifacts in FDK reconstructions.

4.3. Additional study results

4.3.1. Reconstructions with image- ℓ_1 and image- ℓ_2 constraints—We have investigated artifact reduction with the optimization programs in table 1, and show results above obtained from the optimization programs with image-TV constraint. In the discussion below, we present reconstructions obtained from optimization programs with image- ℓ_1 and image- ℓ_2 constraints instead, i.e.

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} D_{\ell_2}(\mathbf{f}) \quad \text{s. t.} \quad \|\mathbf{f}\|_{\ell_1} \leq l_1 \quad \text{and} \quad f_j \geq 0, \quad (15)$$

and

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} D_{\ell_2}(\mathbf{f}) \quad \text{s. t.} \quad \|\mathbf{f}\|_{\ell_2} \leq l_2 \quad \text{and} \quad f_j \geq 0, \quad (16)$$

which are denoted as $D_{\ell_2-\|\mathbf{f}\|_{\ell_1}}$ and $D_{\ell_2-\|\mathbf{f}\|_{\ell_2}}$, respectively, in table 1. For each study, image-constraint parameters l_1 and l_2 were determined as described in section 3.2. Using the CP algorithm to solve the programs, we reconstruct images from the Rando and Catphan data, and display them in figure 7. Reconstructions in figure 7 demonstrate that optimization programs with image- ℓ_1 and image- ℓ_2 constraints appear to be considerably less effective than the optimization programs with image-TV constraint in reducing the FDK-reconstruction artifacts. Moreover, as expected, reconstructions with an image- ℓ_1 constraint appear with noisy, sprinkled-salt textures because both Rando and Catphan are highly non-sparse. It is also interesting to notice that reconstructions based upon program $D_{\ell_2-\|\mathbf{f}\|_{\ell_2}}$ in equation (15) have an appearance resembling that of the corresponding FDK reconstructions. This is because program $D_{\ell_2-\|\mathbf{f}\|_{\ell_2}}$ without image-positivity constrain is equivalent to the Tikhonov-Phillips method, which has been show to be an analogy of the filtered-back projection (i.e. the FDK) algorithm (Natterer and Haderer 1980, Natterer and Wübbeling 2001).

4.3.2. Effect of different weighting matrices on image reconstruction—In

analytic, such as FDK, reconstruction, a weighting function used for designing a weighting matrix is required to satisfy the normalization and continuity conditions in equations (12) and (13) for avoiding yielding significant numerical artifacts. We have investigated weighting functions with different forms in optimization reconstructions. Using a weighting matrix formed from a weighting function

$$\begin{cases} w(u, v, \phi) = \frac{1}{2} & 0 \leq \phi < 2\gamma_m - 2\gamma \\ w(u, v, \phi) = 1 & 2\gamma_m - 2\gamma \leq \phi < \pi - 2\gamma \\ w(u, v, \phi) = \frac{1}{2} & \pi - 2\gamma \leq \phi < \pi + 2\gamma_m \\ w(u, v, \phi) = 0 & \pi + 2\gamma_m \leq \phi < 2\pi, \end{cases} \quad (17)$$

which satisfies the normalization condition in equation (12), but not the continuity condition in equation (13), we carried out reconstruction based upon optimization programs considered, and display example reconstructions in figure 8 obtained by solving program $D_{\ell_2-\|\mathbf{f}\|_{TV}}$ in equation (10) from the Rando and Catphan data. As shown in row 1 of figure 8, optimization-based reconstructions are virtually free of prominent numerical artifacts in FDK reconstructions, resulted from the weighting-function discontinuity in ϕ . Furthermore, we studied reconstructions using an identity matrix, corresponding to a uniform weighting

function without satisfying equation (12), and display them in row 2 of figure 8. Again, optimization-based reconstructions appear free of numerical artifacts in the FDK reconstruction, as a consequence that the uniform weighting function does not satisfy the normalization condition. The results of the investigation indicate weighting matrices designed by use of weighting functions satisfying, or without satisfying, equations (12) and/or (13) appear to have little impact on the performance of optimization-based reconstructions.

5. Discussion

In this work, we have investigated optimization-based reconstructions for possible reduction of image artifacts often observed in FDK reconstruction from short-scan CBCT data. In an attempt to appreciate how different designs of optimization-based reconstruction may impact the artifact reduction, numerous convex optimization programs of potential research and application interest were considered. The investigation is enabled by a primal-dual algorithm that is shown mathematically to solve convex optimization programs, especially for those under consideration. The focus of the study is not on quantitative assessment of an optimization-based reconstruction in a specific, clinically or practically, relevant task; it is tailored, instead, to demonstrating whether optimization-based reconstructions can be designed for reducing artifacts observed in FDK reconstruction. The study confirms that optimization-based reconstruction with appropriate image constraints such as image-TV constraint can reduce FDK-reconstruction artifacts in CBCT with a short-scan configuration. As the results demonstrated, a direct benefit of such an artifact reduction can be an improved contrast level of low-contrast anatomic structures that are obscured otherwise by the artifacts.

The performance of optimization reconstructions in a specific imaging task is likely to be dependent upon the forms and parameters specifying the optimization-programs, as the study results demonstrated. In terms of the reduction of artifacts observed in FDK reconstruction from short-scan CBCT data, visual inspection of study results reveals that reconstructions based upon image-TV-constrained data-discrepancy minimization appear to be capable of reducing the artifacts. Furthermore, although the selection of different weighting matrices has a significant impact on FDK reconstruction, it appears to have insignificant effect on optimization-based reconstructions. Using two phantoms of IGRT CBCT relevance, we show that the image-TV-constrained reconstruction remains effective in reduction of off-middle-plane artifacts in FDK reconstructions from short-scan CBCT data for objects with substantially distinct anatomies.

We considered in the study reconstruction only from short-scan CBCT data. However, the work has implications for reconstruction from data acquired with other scanning configurations in CBCT. It is not uncommon that a scan with an angular range between those of short-scan and full-scan configurations is used in practical applications. The investigation and analysis approach in the work can readily be tailored to investigating appropriate optimization-based reconstruction for possible reduction of artifacts that may appear in the corresponding FDK reconstruction. The study method and results can also be

adopted to investigate optimization-based reconstruction from data collected with an offset detector in CBCT.

The evaluation of the study results was based largely on visual inspection, which serves adequately the focus of the study, which is on whether optimization-based reconstruction can reduce some of the prominent artifacts observed in FDK reconstructions from short-scan CBCT data. However, as the results reveal, different optimization-based reconstructions may behave distinctly differently, depending upon numerous factors, including their design parameters, data conditions, and object properties. More importantly, their performance is likely to be task-specific and thus should be measured with metrics of task-performance utility. Quantitative study of task-specific performance of optimization-based reconstructions can be a topic of theoretical and practical interest for future research.

Acknowledgments

The authors would like to thank S Rose and B Chen for useful discussion. This work was supported in part by the National Institutes of Health (NIH) under grants R01s CA158446, CA182264, and EB018102.

References

- Bian J, Siewerdsen JH, Han X, Sidky EY, Prince JL, Pelizzari CA, Pan X. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. *Phys Med Biol*. 2010; 55:6575–99. [PubMed: 20962368]
- Bian J, Yang K, Boone JM, Han X, Sidky EY, Pan X. Investigation of iterative image reconstruction in low-dose breast CT. *Phys Med Biol*. 2014; 59:2659. [PubMed: 24786683]
- Chambolle A, Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imaging Vis*. 2011; 40:1–26.
- Delaney A, Bresler Y, Sunniva C. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Trans Imaging Process*. 1998; 7:204–21.
- Dennerlein F, Noo F, Schöndube H, Lauritsch G, Hornegger J. A factorization approach for cone-beam reconstruction on a circular short-scan. *IEEE Trans Med Imaging*. 2008; 27:887–96. [PubMed: 18599394]
- De Vos W, Casselman J, Swennen G. Cone-beam computerized tomography (CBCT) imaging of the oral and maxillofacial region: a systematic review of the literature. *Int J Oral Maxillofac Surg*. 2009; 38:609–25.
- Elbakri I, Fessler J. Statistical image reconstruction for polyenergetic x-ray computed tomography. *IEEE Trans Med Imaging*. 2002; 21:89–99. [PubMed: 11929108]
- Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am A*. 1984; 1:612–9.
- Grass M, Köhler T, Proksa R. Angular weighted hybrid cone-beam CT reconstruction for circular trajectories. *Phys Med Biol*. 2001; 46:1595. [PubMed: 11419621]
- Grass M, Koppe R, Klotz E, Proksa R, Kuhn M, Aerts H, De Beek JO, Kemkers R. Three-dimensional reconstruction of high contrast objects using C-arm image intensifier projection data. *Comput Med Imaging Graph*. 1999; 23:311–21. [PubMed: 10634143]
- Grills IS, Hugo G, Kestin LL, Galerani AP, Chao KK, Wloch J, Yan D. Image-guided radiotherapy via daily online cone-beam CT substantially reduces margin requirements for stereotactic lung radiotherapy. *Int J Radiat Oncol*. 2008; 70:1045–56.
- Han X, Pearson E, Pelizzari C, Al-Hallaq H, Sidky EY, Bian J, Pan X. Algorithm-enabled exploration of image-quality potential of cone-beam CT in image-guided radiation therapy. *Phys Med Biol*. 2015; 60:4601. [PubMed: 26020490]
- Hott JS, Deshmukh VR, Klopfenstein JD, Sonntag VK, Dickman CA, Spetzler RF, Papadopoulos SM. Intraoperative Iso-C C-arm navigation in craniospinal surgery: the first 60 cases. *Neurosurgery*. 2004; 54:1131–7. [PubMed: 15113467]

- Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *Int J Radiat Oncol Biol Phys.* 2002; 53:1337–49. [PubMed: 12128137]
- Johnson RH, Hu H, Haworth ST, Cho PS, Dawson CA, Linehan JH. Feldkamp and circle-and-line cone-beam reconstruction for 3D micro-CT of vascular networks. *Phys Med Biol.* 1998; 43:929. [PubMed: 9572516]
- Lauritsch G, Boese J, Wigström L, Kemeth H, Fahrig R. Towards cardiac C-arm computed tomography. *IEEE Trans Med ImagING.* 2006; 25:922–34. [PubMed: 16827492]
- Maaß C, Dennerlein F, Noo F, Kachelrieß M. Comparing short scan CT reconstruction algorithms regarding cone-beam artifact performance. *IEEE Nuclear Science Sym and Medical Imaging Conf Record.* 2010:2188–93.
- Mori S, Endo M, Komatsu S, Kandatsu S, Yashiro T, Baba M. A combination-weighted Feldkamp-based reconstruction algorithm for cone-beam CT. *Phys Med Biol.* 2006; 51:3953. [PubMed: 16885617]
- Natterer F, Haderer K. Efficient implementation of ‘optimal’ algorithms in computerized tomography. *Math Method Appl Sci.* 1980; 2:545–55.
- Natterer, F., Wübbeling, F. *Mathematical Methods in Image Reconstruction.* Philadelphia, PA: Society for Industrial and Applied Mathematics; 2001.
- Nien H, Fessler JA. Fast splitting-based ordered-subsets x-ray CT image reconstruction. *Proc 3rd Int Mtg on image formation in X-ray CT.* 2014:291–4.
- Oldham M, Létourneau D, Watt L, Hugo G, Yan D, Lockman D, Kim LH, Chen PY, Martinez A, Wong JW. Cone-beam-CT guided radiation therapy: a model for on-line application. *Radiother Oncol.* 2005; 75:271–E1. [PubMed: 15890419]
- Orth RC, Wallace MJ, Kuo MD. C-arm cone-beam CT: general principles and technical considerations for use in interventional radiology. *J Vasc Interv Radiol.* 2008; 19:814–20. [PubMed: 18503894]
- Pan X, Sidky EY, Vannier M. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Probl.* 2009; 25:123009.
- Parker DL. Optimal short scan convolution reconstruction for fan beam CT. *Med Phys.* 1982; 9:254–7. [PubMed: 7087912]
- Sidky EY, Jørgensen JH, Pan X. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm. *Phys Med Biol.* 2012; 57:3065. [PubMed: 22538474]
- Sidky EY, Kao KM, Pan X. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *J X-ray Sci Technol.* 2006; 14:119–39.
- Sidky EY, Kraemer DN, Roth EG, Ullberg C, Reiser IS, Pan X. Analysis of iterative region-of-interest image reconstruction for x-ray computed tomography. *J Med Imaging.* 2014; 1:031007.
- Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys Med Biol.* 2008; 53:4777–807. [PubMed: 18701771]
- Siewerdsen JH, Moseley DJ, Burch S, Bisland SK, Bogaards A, Wilson BC, Jaffray DA. Volume CT with a flat-panel detector on a mobile, isocentric C-arm: pre-clinical investigation in guidance of minimally invasive surgery. *Med Phys.* 2005; 32:241–54. [PubMed: 15719975]
- Smitsmans MH, De Bois J, Sonke JJ, Betgen A, Zijp LJ, Jaffray DA, Lebesque JV, Van Herk M. Automatic prostate localization on cone-beam CT scans for high precision image-guided radiotherapy. *Int J Radiat Oncol.* 2005; 63:975–84.
- Stsepankou D, Arns A, Ng S, Zygmanski P, Hesser J. Evaluation of robustness of maximum likelihood cone-beam CT reconstruction with total variation regularization. *Phys Med Biol.* 2012; 57:5955. [PubMed: 22964760]
- Tang J, Nett BE, Chen GH. Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms. *Phys Med Biol.* 2009; 54:5781–804. [PubMed: 19741274]
- Tang X, Hsieh J, Hagiwara A, Nilsen RA, Thibault JB, Drapkin E. A three-dimensional weighted cone beam filtered backprojection (CB-FBP) algorithm for image reconstruction in volumetric CT under a circular source trajectory. *Phys Med Biol.* 2005; 50:3889. [PubMed: 16077234]

- Wallace MJ, Kuo MD, Glaiberman C, Binkert CA, Orth RC, Soulez G. Three-dimensional C-arm cone-beam CT: applications in the interventional suite. *J Vasc Interv Radiol.* 2008; 19:799–813. [PubMed: 18503893]
- Wang AS, Stayman JW, Otake Y, Kleinszig G, Vogt S, Gallia GL, Khanna AJ, Siewerdsen JH. Soft-tissue imaging with C-arm cone-beam CT using statistical reconstruction. *Phys Med Biol.* 2014; 59:1005. [PubMed: 24504126]
- Wiesent K, Barth K, Navab N, Durlak P, Brunner T, Schuetz T, Seissler W. Enhanced 3D-reconstruction algorithm for C-arm systems suitable for interventional procedures. *IEEE Trans Med Imaging.* 2000; 19:391–403. [PubMed: 11021683]

Appendix

The pseudo-codes for the CP algorithm used are given below:

Algorithm

Pseudo-codes of the CP algorithm for solving programs in table 1

INPUT: $\mathbf{g}_m$; $\mathcal{H}$, $\mathcal{W}$, and $\mathcal{X}$; ζ , and λ

- 1: $\nu \leftarrow \nu_{\mathcal{H}}, L \leftarrow \|\mathcal{H}\mathcal{X}\|_{SV}, \tau \leftarrow 1/L, \sigma \leftarrow 1/L$
- 2: $\theta \leftarrow 1, n \leftarrow 0$
- 3: INITIALIZE: $\mathbf{f}_0, \mathbf{p}_0$, and $\mathbf{q}_0$ to zero
- 4: $\tilde{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$
- 5: **repeat**
- 6: $\mathbf{p}_{n+1} \leftarrow \Theta(\lambda, \mathbf{p}_n, \mathcal{W}, \mathcal{H}\tilde{\mathbf{f}}_n, \mathbf{g}_m)$
- 7: $\mathbf{q}_{n+1} \leftarrow \Omega(\mathbf{q}_n, \nu, \sigma, \mathcal{X}\tilde{\mathbf{f}}_n, \zeta)$
- 8: $\mathbf{f}_{n+1} \leftarrow \text{pos}[\mathbf{f}_n - \tau(\mathcal{H}^T\mathcal{W}^T\mathbf{p}_{n+1} + \nu\mathcal{X}^T\mathbf{q}_{n+1})]$
- 9: $\mathbf{f}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$
- 10: $n \leftarrow n + 1$
- 11: **until** the practical convergence conditions are satisfied
- 12: OUTPUT: image $\mathbf{f}_{\text{conv}} \leftarrow \mathbf{f}_n$

In the algorithm, vector $\mathbf{g}_m$ of size $M = M_u \times M_v \times M_\phi$ denotes measured data, where M_u and M_v are number of detector bins along the u - and v -axis on the detector plane, and M_ϕ represents the number of views at which data are collected; vector $\mathbf{f}_n$ of size $N = N_x \times N_y \times N_z$ image at the n th iteration, where N_x, N_y , and N_z are the numbers of voxels along the x -, y -, and z -axis, respectively; $\mathbf{f}_{\text{conv}}$ the convergent reconstruction when the practical convergence conditions are satisfied; $\mathcal{W}$ and $\mathcal{H}$ weighting and system matrices; parameter

$\nu_{\mathcal{H}} = \frac{\|\mathcal{W}\mathcal{H}\|_{SV}}{\|\mathcal{X}\|_{SV}}$, where $\|\cdot\|_{SV}$ denotes the largest singular value of a matrix (Sidky *et al* 2012);

matrix $\mathcal{H}\mathcal{X}$ has a transpose $\mathcal{H}\mathcal{X}^T = (\mathcal{H}^T\mathcal{W}^T, \nu\mathcal{X}^T)$, where the superscript ‘T’ indicates a transpose operation; $\text{pos}(\mathbf{f})$ enforces $f_j = 0$, if $f_j < 0$, where f_j denotes the j th entry of vector $\mathbf{f}$ of size N ; vectors $\mathbf{p}_n$ and Θ are of size M , whereas vectors $\mathbf{q}_n$ and Ω have the same size, which depends upon the programs considered. The specifics of vectors Θ, Ω , along with matrix $\mathcal{X}$, are discussed below for optimization programs under study. Parameter ζ denotes ν_{l_1}, ν_{l_2} , and ν_{t_1} , respectively, for programs with image- l_1 , image- l_2 , and image-TV constraints. For each study, the values of l_1, l_2 , and t_1 were determined in section 3.2.

Algorithm parameter λ has no impact mathematically on the feasible solution set specified

by an optimization program, and it can, however, affect the convergence rate of the algorithm. We select λ between 0.01 to 1.0 in the study, because previous empirical studies have shown that this selection generally yields a reasonable convergence rate.

For the convenience of discussing the determination of Θ and Ω below, we consider two vectors $\mathbf{a}$ and $\mathbf{b}$ of sizes kN and N , where $k \geq 1$ is a positive integer. Let a_j denote the j th entry of $\mathbf{a}$, $j = 1, 2, \dots, kN$, and b_j the j th entry of $\mathbf{b}$, $j = 1, 2, \dots, N$, respectively. We use vectors $\mathbf{ab}$ and $\mathbf{a/b}$ of sizes kN to denote the multiplication and division of $\mathbf{a}$ and $\mathbf{b}$, with entry j given by

$$\begin{aligned} (\mathbf{ab})_j &= a_j b_{j_m} \\ (\mathbf{a/b})_j &= a_j / b_{j_m}, \end{aligned}$$

where $j = 1, 2, \dots, kN$, and $j_m = \text{mod}(j, N)$ indicates the remainder of j dividing by N , and $b_{j_m} \neq 0$. In particular, when $k = 1$, $\mathbf{a}$ and $\mathbf{b}$ are of the same size N , and vectors $\mathbf{ab}$ and $\mathbf{a/b}$ thus are of size N and depict element-wise multiplication and division.

A. 1. Updates In line 6

Update vector $\Theta(\lambda, \mathbf{p}_n, \mathcal{W}, \mathcal{H}\bar{\mathbf{f}}_n, \mathbf{g}_m)$ of size M in line 6 depends upon the specific form of data fidelity considered. For programs in column 2 of table 1,

$$\Theta(\lambda, \mathbf{p}_n, \mathcal{W}, \mathcal{H}\bar{\mathbf{f}}_n, \mathbf{g}_m) = [\mathbf{p}_n + \sigma \mathcal{W}(\mathcal{H}\bar{\mathbf{f}}_n - \mathbf{g}_m)] / (1 + \sigma / \lambda); \quad (\text{A.1})$$

for optimization programs in column 3 of table 1,

$$\Theta(\lambda, \mathbf{p}_n, \mathcal{W}, \mathcal{H}\bar{\mathbf{f}}_n, \mathbf{g}_m) = \frac{\mathbf{p}_n + \sigma \mathcal{W}(\mathcal{H}\bar{\mathbf{f}}_n - \mathbf{g}_m)}{\max(\mathbf{1}_D, |\mathbf{p}_n + \sigma \mathcal{W}(\mathcal{H}\bar{\mathbf{f}}_n - \mathbf{g}_m)| / \lambda)}, \quad (\text{A.2})$$

where $\mathbf{1}_D$ denotes a vector of size M with all entries set to 1, $|\mathbf{p}|$ takes entry-wise absolute values of vector $\mathbf{p}$ of size M , and $\max(\cdot)$ is performed component-wise; and for programs in column 4 in table 1,

$$\Theta(\lambda, \mathbf{p}_n, \mathcal{W}, \mathcal{H}\bar{\mathbf{f}}_n, \mathbf{g}_m) = \frac{1}{2} [\lambda_D + \mathbf{p}_n + \sigma \mathcal{W} \mathcal{H}\bar{\mathbf{f}}_n] - \sqrt{[\lambda_D - \mathbf{p}_n - \sigma \mathcal{W} \mathcal{H}\bar{\mathbf{f}}_n]^2 + 4\sigma \lambda \mathcal{W} \mathbf{g}_m}, \quad (\text{A.3})$$

where λ_D is a vector of size M with entries set to λ , and $\sqrt{\cdot}$ takes entry-wise square roots of vector $\mathbf{p}$ of size M .

A.2. Updates in line 7

For programs in row 2, $\mathcal{X}$ is an identity matrix of size N , and vector $\mathbf{\Omega}$ of size N is given by

$$\mathbf{\Omega}(\mathbf{q}_n, \nu, \sigma, \mathcal{X}\bar{\mathbf{f}}_n, \zeta) = \mathbf{0}, \quad (\text{A.4})$$

where $\mathbf{0}$ is a vector of size N with entries set to zero.

For program in rows 3 and 4 of table 1, $\mathcal{X}$ is an identity matrix of size N , image-constraint parameter $\zeta = \nu l_1$ for programs in row 3 and $\zeta = \nu l_2$ for programs in row 4, and vector $\mathbf{\Omega}$ of size N is given by

$$\mathbf{\Omega}(\mathbf{q}_n, \nu, \sigma, \mathcal{X}\bar{\mathbf{f}}_n, \zeta) = \mathbf{\varrho}(\mathbf{1}_I - \sigma \mathbf{s} / |\mathbf{\varrho}|_{\text{MAG}}), \quad (\text{A.5})$$

where vector $\mathbf{\varrho}$ of size N is defined as

$$\mathbf{\varrho} = \mathbf{q}_n + \sigma \nu \bar{\mathbf{f}}_n,$$

vector $\mathbf{1}_I$ of size N with entries set to 1, and vector $\mathbf{s}$ of size N is given by

$$\mathbf{s} = \text{ProjOnto}_{\ell_p \text{Ball}_c}(|\mathbf{\varrho}|_{\text{MAG}} / \sigma), \quad (\text{A.6})$$

operator $\text{ProjOnto}_{\ell_p \text{Ball}_c}$ projects vector $|\mathbf{\varrho}|_{\text{MAG}} / \sigma$ onto the ℓ_p -ball of scale c , $|\mathbf{\varrho}|_{\text{MAG}}$ depicts a vector of size N with entry j given by $(|\mathbf{\varrho}|_{\text{MAG}})_j = |\mathbf{\varrho}_j|$, and $\mathbf{\varrho}_j$ indicates the j th entry of $\mathbf{\varrho}$. Furthermore, $p = 1$ and $c = \nu l_1$ for programs in row 3 of table 1, and $p = 2$ and $c = \nu l_2$ for programs in row 4 of table 1.

For program in row 5 of table 1, matrix $\mathcal{X} = \nabla$ denotes a spatial gradient matrix of size $3N \times N$, with its transpose $\nabla^\top = (\nabla_x^\top, \nabla_y^\top, \nabla_z^\top)$, where matrices ∇_x , ∇_y , and ∇_z of size $N \times N$ represent the finite difference matrices along x -, y -, and z -axis, respectively, yielding vectors $\nabla_x \mathbf{f}$, $\nabla_y \mathbf{f}$, and $\nabla_z \mathbf{f}$ of size N , which are used to form vector $\nabla \mathbf{f}$ of size $3N$ in a concatenated form in the order of x , y , and z , $|\nabla \mathbf{f}|_{\text{MAG}}$ depicts a vector of size $3N$ with entry j given by $(|\nabla \mathbf{f}|_{\text{MAG}})_j = \sqrt{(\nabla \mathbf{f})_j^2 + (\nabla \mathbf{f})_{j+N}^2 + (\nabla \mathbf{f})_{j+2N}^2}$, and $(\nabla \mathbf{f})_j$ indicates the j th entry of vector $\nabla \mathbf{f}$. Moreover, image-constraint parameter $\zeta = \nu l_1$, and vector $\mathbf{\Omega}$ of size $3N$ is given by

$$\mathbf{\Omega}(\mathbf{q}_n, \nu, \sigma, \mathcal{X}\bar{\mathbf{f}}_n, \zeta) = \mathbf{\varrho}(\mathbf{1}_I - \sigma \mathbf{s} / |\mathbf{\varrho}|_{\text{MAG}}), \quad (\text{A.7})$$

where vector $\mathbf{\varrho}$ of size $3N$ is defined as

$$\boldsymbol{\varrho} = \mathbf{q}_n + \sigma \nu \nabla \bar{\mathbf{f}}_n,$$

vector $\mathbf{1}_I$ is of size N with entries set to 1, and vector $\mathbf{s}$ of size N is given by

$$\mathbf{s} = \text{ProjOnto}_{\ell_1} \text{Ball}_{\nu t_1}(|\boldsymbol{\varrho}|_{\text{MAG}}/\sigma), \quad (\text{A.8})$$

operator **ProjOntoBall** $_{\nu t_1}$ projects vector $|\boldsymbol{\varrho}|_{\text{MAG}}/\sigma$ onto the ℓ_1 -ball of scale νt_1 (Sidky *et al* 2014), $|\boldsymbol{\varrho}|_{\text{MAG}}$ depicts a vector of size N with entry j given by

$$(|\boldsymbol{\varrho}|_{\text{MAG}})_j = \sqrt{\varrho_j^2 + \varrho_{j+N}^2 + \varrho_{j+2N}^2}, \text{ and } \varrho_j \text{ indicates the } j\text{th entry of vector } \boldsymbol{\varrho}.$$

Figure 1.

ROIs indicated by rectangles within transverse slices at 3.37 cm (column 1) and 7.32 cm (column 2) from the middle plane of the Rando phantom, and at 4.15 cm (column 3) and 6.0 cm (column 4) from the middle plane of the Catphan phantom. ROI images are displayed in zoomed-in views below.

Figure 2.

Reconstructions within the middle plane of the Rando phantom obtained using the FDK algorithm, and by using the CP algorithm solving, respectively, optimization programs shown in table 1. Display window: $[0.22, 0.30] \text{ cm}^{-1}$.

Figure 3.

Reconstructions, together with zoomed-in images within ROIs defined in figure 1, of the Rando phantom within transverse slices at 3.37 cm (row 1) and 7.32 cm (row 2) from the middle plane, reconstructed by use of the FDK algorithm and the CP algorithm solving, respectively, optimization programs in row 5 of table 1. Arrows in the reconstruction in panel 4 of row 2 highlight the low-contrast structures in the Rando phantom. Display window: $[0.22, 0.30] \text{ cm}^{-1}$.

Figure 4.

Reconstructions, together with zoomed-in images within ROIs defined in figure 1, of the Catphan phantom within transverse slices at 4.15 cm (row 1) and 6.00 cm (row 2) from the middle plane, reconstructed by use of the FDK algorithm and the CP algorithm solving, respectively, optimization programs in row 5 of table 1. Display window: $[0.22, 0.32] \text{ cm}^{-1}$ for row 1 and $[0.22, 0.35] \text{ cm}^{-1}$ for row 2.

Figure 5.

Data- ℓ_2 (left), image-TV (middle), and conditional primal-dual gap (cPD) (right), as functions of iteration number n for Rando-phantom reconstruction by using the CP algorithm solving program $D_{\ell_2}-\|f\|_{TV}$ in equation (10). Dashed line in image-TV plot (middle) is specified by image-constraint parameter t_1 determined. The plot scales are in arbitrary units.

Figure 6.

Reconstructions of the Rando phantom (row 1) within transverse slice at 7.32 cm (row 1), and of Catphan (row 2) phantom within transverse slice at 6.0 cm, from their respective middle planes at iteration numbers indicated, along with their respective, convergent reconstructions (column 4), obtained with the CP algorithm solving program $D_{\ell_2} - \|\cdot\|_{TV}$ in equation (10). Display windows $[0.22, 0.30] \text{ cm}^{-1}$ and $[0.22, 0.35] \text{ cm}^{-1}$ are for Rando and Catphan images, respectively.

Figure 7.

Reconstructions of the Rando phantom (row 1) within transverse slice at 7.32 cm, and of the Catphan phantom (row 2) within transverse slice at 6.0 cm, from their respective middle planes, reconstructed by use of the FDK algorithm, and by using the CP algorithm solving optimization programs in column 3 of table 1, respectively. Display windows $[0.22, 0.30] \text{ cm}^{-1}$ and $[0.22, 0.35] \text{ cm}^{-1}$ are for the Rando and Catphan images, respectively.

Figure 8.

Images of the Rando phantom within transverse slice at 7.32 cm, and of the Catphan phantom within transverse slice at 6.0cm, from their respective middle planes, reconstructed by using the FDK algorithm (column 1 & 3), and by using the CP algorithm (column 2 & 4) solving optimization program $D_{\ell_2}||f||_{TV}$ in equation (10) with weighting matrices calculated from a non-smooth weighting function in equation (17) (row 1) and a smooth, but uniform, weighting function (row 2), respectively. Display windows $[0.22, 0.35] \text{ cm}^{-1}$ and $[0.22, 0.30] \text{ cm}^{-1}$ are for the Rando and Catphan images, respectively.

Table 1

Symbols of optimization programs formed by use of data fidelities in equations (3)–(5) and image constraints in equations (6)–(8) to replace Φ and Ψ in equation (2).

Φ	$D_{\mathbf{f}}(\mathbf{f})$	$D_{\mathbf{g}}(\mathbf{f})$	$D_{\text{KL}}(\mathbf{f})$	α
Ψ				
N/A	$D_{\mathbf{f}_1}$	$D_{\mathbf{f}_2}$	D_{KL}	N/A
$\ \mathbf{f}\ _{l_1}$	$D_{\mathbf{f}_1-\ \mathbf{f}\ _{l_1}}$	$D_{\mathbf{f}_2-\ \mathbf{f}\ _{l_1}}$	$D_{\text{KL}-\ \mathbf{f}\ _{l_1}}$	l_1
$\ \mathbf{f}\ _{l_2}$	$D_{\mathbf{f}_1-\ \mathbf{f}\ _{l_2}}$	$D_{\mathbf{f}_2-\ \mathbf{f}\ _{l_2}}$	$D_{\text{KL}-\ \mathbf{f}\ _{l_2}}$	l_2
$\ \mathbf{f}\ _{\text{TV}}$	$D_{\mathbf{f}_1-\ \mathbf{f}\ _{\text{TV}}}$	$D_{\mathbf{f}_2-\ \mathbf{f}\ _{\text{TV}}}$	$D_{\text{KL}-\ \mathbf{f}\ _{\text{TV}}}$	t_1

Note: For example, using $D_{\text{KL}}(\mathbf{f})$ and $\|\mathbf{f}\|_{\text{TV}}$ to replace Φ and Ψ in equation (2) yields program $D_{\text{KL}-\|\mathbf{f}\|_{\text{TV}}}$, as shown in equation (11) below.