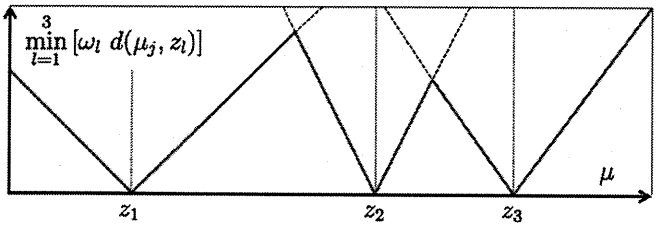

Figure 1. The proposed penalty function in equation (8) corresponding to $L = 3$.

The most advantageous feature, of the proposed penalty function in equation (8), is that the required *a priori* information is limited to only a small number of intensity values, which can be easily available in several clinical applications. An illustration of the proposed penalty function is shown in figure 1.

3. Methods

3.1. Reconstruction algorithm

The main difficulty in minimizing the objective function in equation (6) is that the proposed penalty term in equation (8) is based on the ℓ_1 norm, which is neither convex nor differentiable. Therefore, it is not possible to use the ordinary gradient-type iterative method to optimize the objective function. Alternatively, the well-known majorization–minimization (MM) strategy (Fessler and Hero 1995, Daubechies *et al* 2004) is used to derive a separable quadratic function. At each iteration k, the non-separable part of the main objective function in equation (6) is approximated by a separable quadratic function $\tilde{L}(\mu; \mu^k)$ around $\mu = \mu^k$ such that

$$\tilde{L}(\mu^k; \mu^k) = L(\mu^k), \tag{10}$$

$$\partial\tilde{L}(\mu; \mu^k)/\partial\mu_j|_{\mu=\mu^k} = \partial L(\mu)/\partial\mu_j|_{\mu=\mu^k}. \tag{11}$$

The approximate quadratic majorization to $L(\mu)$ is formulated as

$$\tilde{L}(\mu; \mu^k) = L(\mu^k) + \nabla L(\mu)|_{\mu=\mu^k}(\mu - \mu^k)^\top + \frac{1}{2}Q(\mu^k)(\mu_j - \mu_j^k)^2 \tag{12}$$

$$= \frac{1}{2}\sum_{j=1}^{n}\frac{\sum_{i=1}^{m} a_{ij}\langle a_i, \mu^k\rangle b_i \exp(-\langle a_i, \mu^k\rangle)}{\mu_j^k}(\mu_j - p_j)^2 + T(\mu^k), \tag{13}$$

where $T(\mu^k)$ is the term independent of μ,

$$Q(\mu^k) = \sum_{j=1}^{n}\frac{\sum_{i=1}^{m} a_{ij}\langle a_i, \mu^k\rangle b_i \exp(-\langle a_i, \mu^k\rangle)}{\mu_j^k}, \tag{14}$$

and

$$p_j = \mu_j^k + \mu_j^k\frac{\sum_{i=1}^{m} a_{ij}(b_i \exp(-\langle a_i, \mu^k\rangle) - y_i)}{\sum_{i=1}^{m} a_{ij}\langle a_i, \mu^k\rangle b_i \exp(-\langle a_i, \mu^k\rangle)}. \tag{15}$$

The resulting separable objective function $\tilde{f}_\beta(\mu; \mu^k) = \tilde{L}(\mu; \mu^k) + \beta U(\mu)$ is minimized analytically to obtain the next iterate μ^{k+1}. This approximation is useful to obtain a separable objective function when the penalty term is non-differentiable such as the case of ℓ_1 norm

distance. We consider the proposed penalty term of intensity information in equation (8), in which the separable objective function $\tilde{f}_\beta(\mu;\mu^k)$ can be written as

$$\tilde{f}_\beta(\mu;\mu^k) = \sum_{j=1}^{n} \beta\omega_h[c_{j,h}(\mu_j - p_j)^2 + d(\mu_j, z_h)] + T(\mu^k), \tag{16}$$

$$h = \{h \in \{1,\ldots,L\} : d(\mu_j, z_h) = \min_{l=1}^{L} d(\mu_j, z_l)\},$$

$$c_{j,h} = \frac{1}{2\beta\omega_h\mu_j^k} \sum_{i=1}^{m} a_{ij}\langle a_i, \mu^k\rangle b_i \exp(-\langle a_i, \mu^k\rangle). \tag{17}$$

The proposed algorithm, which we called intensity-based MAP (iMAP), for transmission tomography is implemented through the following steps.

(i) (Initialization step) Estimate the intensity prior $z = (z_1,\ldots,z_L)$, set the initial image μ^0 to a uniform positive value and set the iteration number $k \to 0$.

(ii) (Majorization step) The objective function $f(\mu)$ is approximately majorized around the current estimate μ^k using equation (13) to the separable function in equation (16).

(iii) (Minimization) The separable objective function $\tilde{f}(\mu;\mu^k)$ is minimized over $\mu \geqslant 0$ to obtain the image estimate μ^{k+1} in the next iterate:

$$q = \arg\min_{\mu\geqslant 0} \tilde{f}_\beta(\mu;\mu^k), \quad \mu_j^{k+1} = \max(q_j, \epsilon), \tag{18}$$

where ϵ is a small value to guarantee that $\mu_j^{k+1} > 0$.

(iv) (Iterate condition) Set iteration number $k \to k+1$, and repeat steps (i)–(iii) until reaching a stopping criterion.

The main theoretical framework for the derivation of the iMAP algorithm is the same as that presented in the earlier work for the anatomical-MAP-EM method for PET/SPECT imaging (Mameuda and Kudo 2007). However, the present method investigates a different imaging problem in transmission CT, where the essential target of the previous work was to reconstruct PET/SPECT images with better noise/contrast properties. Moreover, the intensity prior, which is more appropriate for the transmission CT imaging, is an original contribution of this paper.

It is known that in the generic ℓ_1 norm prior, the convergence of the iterative algorithm is guaranteed (Daubechies *et al* 2004, Loris 2009). In this paper, we have used the MM strategy to generate an approximate separable function for which the convergence properties of the proposed algorithm are not clear. However, the same framework can be used with the exact surrogate function presented by Erdŏgan and Fessler (1999). In this case, the algorithm monotonically decreases the objective function toward a local minimum. Nevertheless, as the objective function is non-convex, there is no proof that the algorithm will achieve the global convergence. Recently, it has become interesting to study the performance of similar algorithms and some papers have discussed the convergence properties of iterative hard thresholding (e.g. Blumensath and Davies 2010).

3.2. Intensity multi-thresholding

The minimization in equation (18) can now be implemented directly. It is easy to prove that the optimization of the following simple separable objective function:

$$x^* = \arg\min_x \frac{\alpha}{2}(x-a)^2 + \beta|x-b|, \tag{19}$$

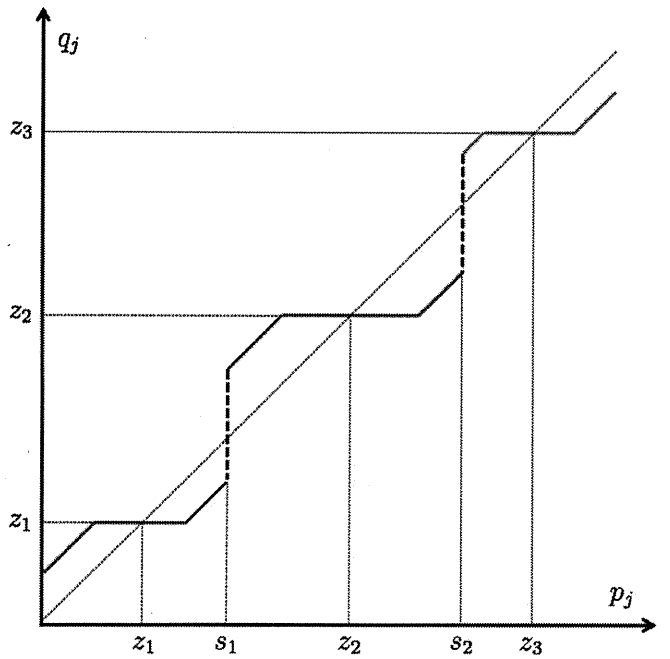

Figure 2. The multi-thresholding function in equation (21) with $L = 3$.

can be found through the following simple soft-thresholding function:

$$x^* = \begin{cases} a + \beta/\alpha & (a < b - \beta/\alpha) \\ b & (b - \beta/\alpha \leqslant a \leqslant b + \beta/\alpha) \\ a - \beta/\alpha & (a > b + \beta/\alpha). \end{cases} \tag{20}$$

In the same way, it is direct to estimate a multi-thresholding function corresponding to the minimization of the objective function in equation (16). Therefore, the iterative reconstruction algorithm is implemented through the following simple procedure. First, we compute the term p_j in equation (15) which is the same as a single iteration of the Convex algorithm (Lange and Fessler 1995). Then, the minimization is implemented through the following multi-thresholding function:

$$q_j = \begin{cases} p_j + 1/(2c_{j,l}) & (s_{l-1} < p_j < z^-_{j,l}) \\ z_l & (\max(z^-_{j,l}, s_{l-1}) \leqslant p_j \leqslant \min(z^+_{j,l}, s_l)) \\ p_j - 1/(2c_{j,l}) & (z^+_{j,l} < p_j \leqslant s_l) \end{cases}, \quad (l = 1, \ldots, L). \tag{21}$$

with $z^+_{j,l} = z_l + 1/(2c_{j,l})$, $z^-_{j,l} = z_l - 1/(2c_{j,l})$, $s_l = (\omega_l z_l + \omega_{l+1} z_{l+1})/(\omega_l + \omega_{l+1})$, $s_0 = -\infty$ and $s_L = \infty$.

This multi-thresholding function can be expressed as a combination of multiple successive soft-thresholding functions; each is implemented around a single value of the *a priori* known intensity values z_l, $l = 1, \ldots, L$. The practical interpretation of this thresholding operation is as follows. If the computed pixel value p_j is closed to the intensity value z_l (i.e. located inside the thresholding window controlled by the parameter ω_l), the pixel value is trimmed to z_l. Otherwise, the pixel value p_j is shifted *softly* toward the closest value of z_l. An illustration of the multi-thresholding functions is in figure 2.

In MAP reconstruction algorithms, the hyper-parameter β in equation (6) is known as the regularization parameter which handles the strength of the prior term. In the iMAP algorithm, β controls the size of the thresholding window in the multi-thresholding function in (21). Obviously, the value of β should be carefully selected. On the one hand, if β is assigned to a rather large value, some low-frequency details in the reconstructed image will be lost due to the strong thresholding. On the other hand, if β is relatively small, the merit of thresholding is weak and the proposed method behaves similar to the conventional reconstruction algorithm. We propose to use a dynamic value for this parameter by starting with a relatively large value and then gradually decrease it according to the following rule:

$$\lim_{k\to\infty} \beta_k = \delta \qquad \text{and} \qquad \sum_{k=0}^{\infty} \beta_k = \infty, \tag{22}$$

where δ is a small value empirically defined to represent the noise magnitude and equals to zero for noise-free data. In the experimental studies presented in this paper, we have used the ordered subset version of the proposed algorithm (OS-iMAP) to speed up the reconstruction, and the under-relaxation parameter is implemented using the following simple rule:

$$\beta_k = (k_{\max} + 1)\beta/(k + 1), \tag{23}$$

where $k_{\max}$ is the number of iterations used in the reconstruction. The use of this dynamic hyper-parameter is useful to speed up the convergence in the practical implementation. This approach is based on the observation that, in the early iterations, a strong thresholding is required to remove acute streak artifacts, and a weak thresholding is preferred later to retrieve missing small-size and low-contrast image details. Moreover, it is known that the use of the under-relaxation parameter improves the convergence properties of ordered subset reconstruction by removing the limit cycle behavior (Kole 2005).

The proposed iMAP algorithm has various advantages. First, the penalty function for intensity prior in equation (8) can be easily embedded in several famous transmission SR algorithms, for example, the EM algorithm (Lange and Carson 1984), gradient algorithm (Lange *et al* 1987) and its ordered subset version (Beekman and Kamphuis 2001). Second, the computation cost is almost same as the original method since the cost of the multi-thresholding step can be ignored relative to the computation of a single iteration of the original reconstruction algorithm. Third, the required intensity prior, as we will present in the next section, can be easily estimated with high accuracy in many imaging situations.

3.3. Estimation of intensity prior

In the iMAP algorithm presented above, we assume that a set of intensity values z is *a priori* known. In this section, we discuss possible approaches to compute the intensity prior.

3.3.1. Application-based estimation. In many x-ray CT applications, the anatomical structure as well as the attenuation coefficient of the target object can be estimated with high accuracy. The knowledge of the tube current energy and the exposure time used in the imaging equipment can provide an accurate limited range for the expected intensity values corresponding to the anatomy of the scanned object. For example, the standard intensity values, in Hounsfield units (HU), expected from the average radiation dose used in clinical CT scanners are shown in table 1 (Ohnesorge *et al* 2006). These values can be simply used as intensity prior in several CT imaging applications with weighting values ω that approximately represent the power of each intensity value.

Table 1. Standard attenuation values of body structures and other objects, measured in Hounsfield units (HU).

Anatomical structure	HU range
Air	−1000
Lung	−900 : −800
Fatty tissue	−100 : −50
Water	0
Blood	30 : 50
Muscles	30 : 80
Contrast enhanced blood	200 : 500
Calcification	130 : 500
Bone	500 : 1500
Metal	>1000

3.3.2. Image-based estimation. In some CT applications, such as IGRT, the patient is subject to multiple scans during a short time of treatment, which significantly increase the patient dose to a risky level. A possible alternative imaging protocol that minimizes the dose and also reduces the scan time can be as follows. First, the patient is imaged with a normal dose for a single time. The reconstructed image is used to estimate the intensity values that represent different regions of interest. The later periodic scans are then implemented through a low-dose or small views scan, which can be employed using intensity prior with the proposed method. Furthermore, it is possible to estimate the intensity values from a segmented image obtained from an earlier scan of the same patient or another patient. Moreover, in multi-slice CT imaging, the patient dose can be modulated by using a hybrid scan, which alternates the conventional scan in a single rotation and the low-dose scan in a few later ones. Intensity values obtained from slices reconstructed from normal dose data can be used for the reconstruction of the neighbor slices using data acquired from low-dose scans.

3.3.3. Data-based estimation.. If the *a priori* information about the internal anatomy of the scanned object is unknown, the intensity values can be estimated directly from the projection data through a clustering of image estimate. As shown later in the experimental studies, it is possible to estimate intensity values during the image reconstruction.

4. Experimental results

In this section, the proposed method is evaluated. The experiments are implemented using a simulated phantom and real data obtained from different x-ray CT imaging applications.

4.1. Assessment of image quality

Throughout the experimental data presented here, the image quality is measured using several methods. Along with each reconstruction result, the corresponding true object in simulated data or the FBP reconstruction from complete projection data is illustrated for visual evaluation. We also calculate the root mean square error (RMSE) for image quality evaluation:

$$\mathrm{RMSE}(\mu) = \sqrt{\sum_{j=1}^{n}(\mu_j - \bar{\mu}_j)^2 \Big/ \sum_{j=1}^{n}\bar{\mu}_j}, \tag{24}$$

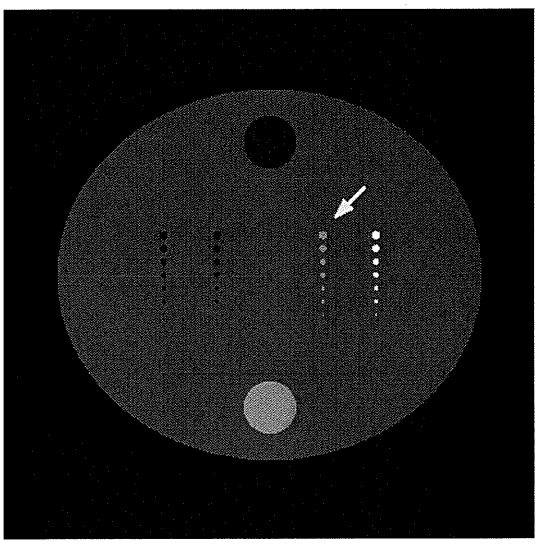

Figure 3. Digital phantom used in the simulation studies. The arrow is pointing to the column of seven resolution inserts used to compute the image contrast.

where $\bar{\mu}$ represents the true object or the FBP image reconstructed from complete data. We also measure the contrast of the reconstructed image using the following formula:

$$\text{contrast}(\mu) = \frac{|\mu_s - \mu_b|}{\mu_s + \mu_b}, \tag{25}$$

where μ_s and μ_b are the average intensity values of selected resolution insert pixels and background pixels, respectively.

4.2. Simulation studies

In the first simulation, we have used a digital phantom shown in figure 3. The phantom size is 100 mm consisting of a uniform background with resolution inserts of different size and contrast. The size of the inserts is ranged from 0.4 to 1.6 mm arranged in four columns each of seven inserts. Two additional inserts of 10 mm are also added to the phantom. The parameters of this phantom are shown in table 2. The image grid was set to 500 × 500 pixels, and the noise-free projection data were measured using a 500 bin detector over 180° with parallel-beam geometry and simple line-length model. The projection data were measured using 20, 10 and 7 projection views and the image reconstruction was implemented using FBP, OS-Convex, ART-TV (Sidky *et al* 2006) and OS-iMAP methods. The image reconstruction using the ART-TV method was implemented through a data enforcement step using the well-known ART algorithm followed by a TV-minimization step in an alternating manner. The parameters of the ART-TV method were optimized to obtain the best image on the RMSE scale. For iterative methods, the iteration number was unified to 100 iterations with an initial uniform image computed from projection data. The number of subsets used for the OS-Convex and OS-iMAP methods was set to 5, 5 and 7 subsets for reconstruction from 20, 10 and 7 view angles, respectively.

The intensity prior for the OS-iMAP method was selected to be the intensity values corresponding to air (region outside the object) and the uniform background of the phantom.

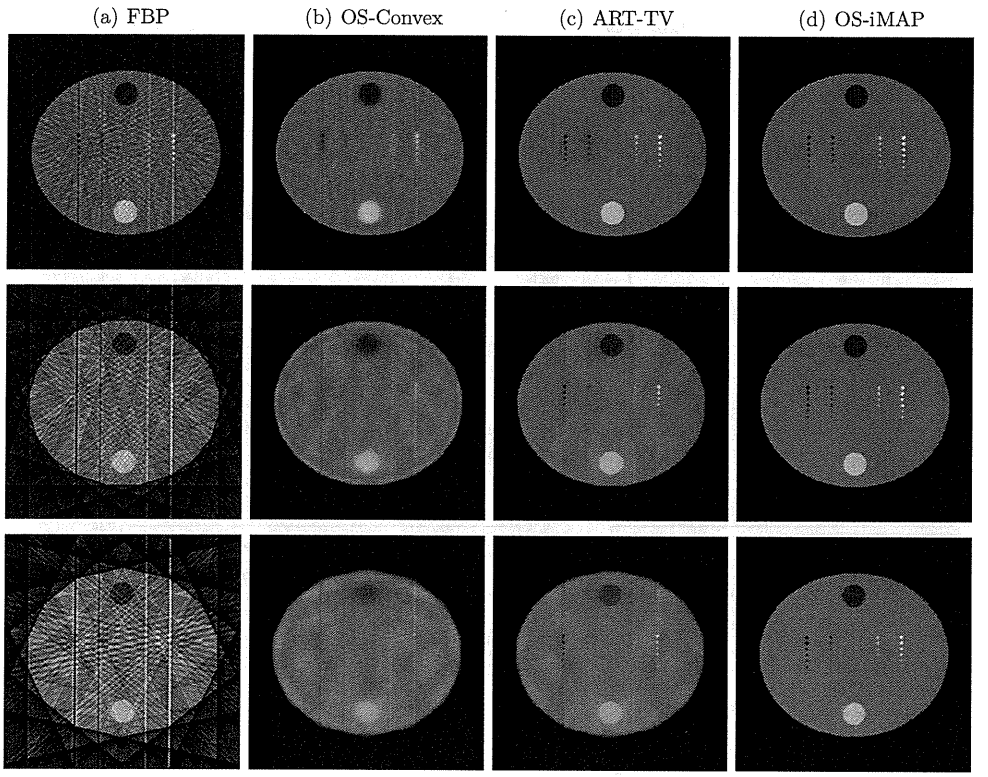

Figure 4. Images reconstructed using (a) standard FBP, (b) OS-Convex, (c) ART-TV and (d) OS-iMAP methods from 20, 10 and 7 projections (from top to bottom).

Table 2. Parameters of a digital phantom shown in figure 3. Objects from numbers 4–10 are repeated with $a = (-0.4, -0.2, 0.2, 0.4)$ and $b = (0.0, 0.5, 1.5, 2.0)$, respectively.

Index	Center coordinates	Major axis	Minor axis	Density (cm^{-1})
1	(0.0, 0.0)	0.8	0.7	1.0
2	(0.0, 0.5)	0.1	0.1	0.5
3	(0.0, −0.5)	0.1	0.1	1.5
4	(a, −0.15)	0.016	0.016	b
5	(a, −0.10)	0.014	0.014	b
6	(a, −0.05)	0.012	0.012	b
7	(a, 0.00)	0.010	0.010	b
8	(a, 0.05)	0.008	0.008	b
9	(a, 0.10)	0.006	0.006	b
10	(a, 0.15)	0.004	0.004	b

We assumed that the inserts are unknown structures. This assumption would probably fit with many diagnostic CT applications, where the intensity value of a uniform background is known while lesions are commonly unknown. The reconstruction was implemented with $L = 2$, $\beta = 0.008$, $z = (0.0, 1.0)$ cm^{-1} and $\omega = (0.01, 0.06)$. From only seven projections, the detectability of the inserts is significantly improved by using the proposed method. This can be confirmed by the visibility of the small-size inserts shown in figure 4 and the corresponding

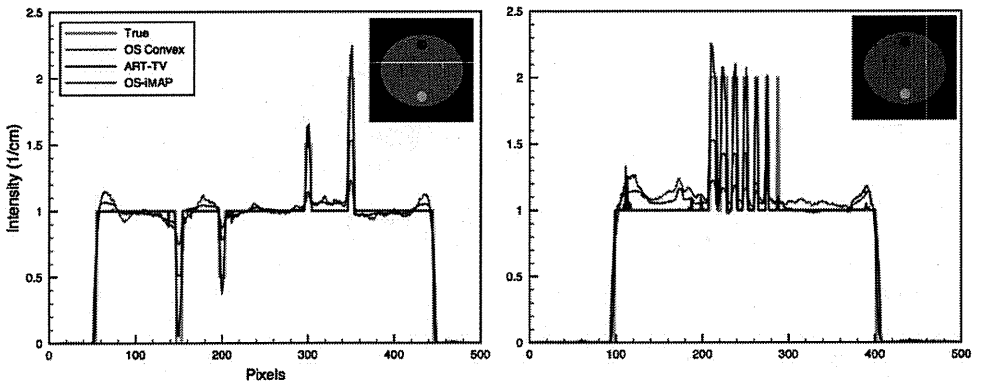

Figure 5. Horizontal and vertical profile plots across the inserts of some results presented in figure 4. Comparison is between images reconstructed from seven projections.

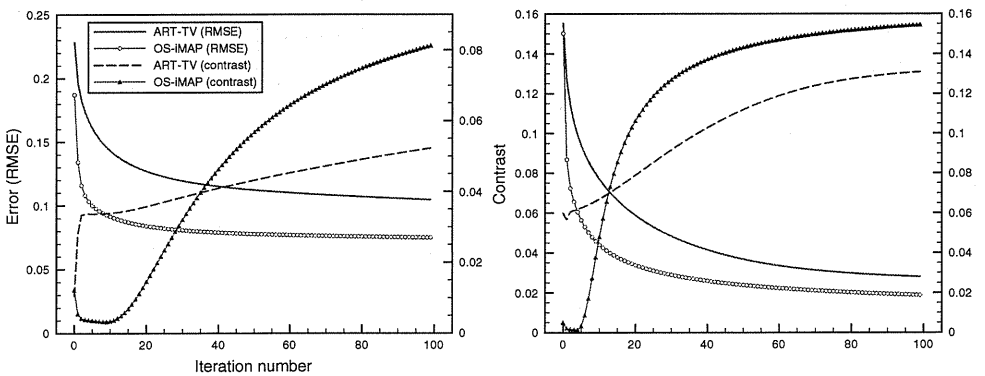

Figure 6. Measured contrast (right vertical axis) of seven inserts, marked by the arrow in figure 3, and RMSE (left vertical axis) for images reconstructed using ART-TV and OS-iMAP methods from 7 (left) and 20 (right) projections.

profile plots in figure 5. Unlike the ART-TV method, it is observed that the large size inserts at 12 and 6 o'clock positions of the phantom still suffer from weak streak artifacts by using the proposed method. This behavior is expected and the reasons are that (1) we did not include any intensity prior corresponding to the inserts and (2) the objective function does not include any uniformity enforcing term (e.g. smoothing penalty). Including smoothing penalty, however, would probably profit image uniformity in the cost of contrast especially for the small-size inserts.

The RMSE and the contrast of the seven inserts, marked by the arrow in figure 3, were computed for the images reconstructed from ART-TV and OS-iMAP methods. Results obtained from 7 and 20 projections are shown in figure 6. From these results, it can be observed that the proposed method achieves a significant improvement in contrast with an acceptable RMSE value, especially when the number of projections is highly under-sampled. The true contrast of the seven inserts is 0.2. The same experiment was repeated with added Poisson noise corresponding to 10^5 photon counts and reconstructed images from 25 projections are shown in figure 7.

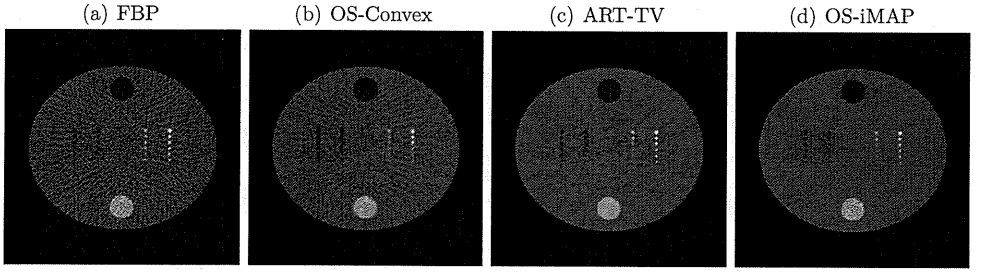

Figure 7. Images reconstructed from 25 projections with added Poisson noise.

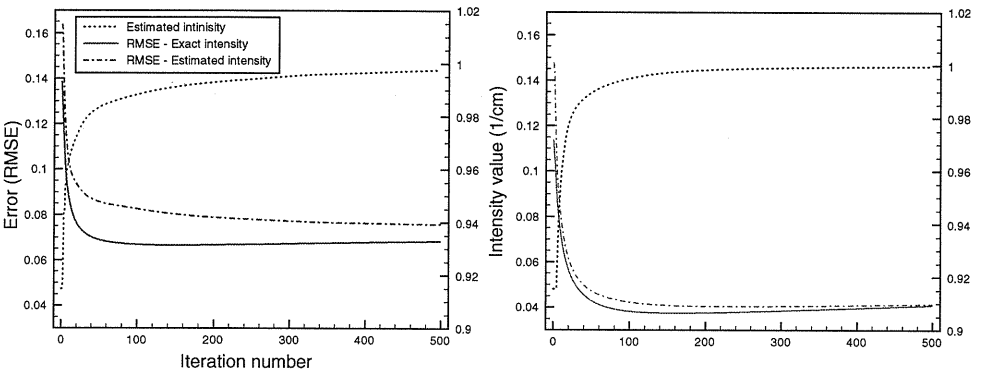

Figure 8. Measured RMSE (left vertical axis) and estimated intensity value (right vertical axis) of the uniform background computed from the OS-iMAP reconstruction from 7 (left) and 20 (right) projections.

In another experiment, we investigate the proposed method when the intensity prior is unavailable and estimated during the reconstruction. The intensity prior were estimated initially from the FBP image and then updated at every iteration using the current image estimate. For simplicity, we limited the estimated intensity prior to only the value corresponding to the uniform background. This value was computed as the median intensity value of the current image estimate. We have used the same experiment settings used in the previous study. The estimated intensity value corresponding to the uniform background and the RMSE of the reconstructed image were calculated for the case of reconstruction from 7 and 20 projections along 500 iterations, and the results are shown in figure 8. These results indicate that the estimated intensity value is approaching to the accurate one as the iteration proceeds. However, more iterations were required to reach the exact intensity value and, therefore, the same image quality.

The effect of the under-relaxation parameter introduced in (22) was evaluated in another experimental study. A noise-free data measured over 20 projection views were used for a reconstruction using different values of the hyper-parameter β. We have used static values of 0.1, 0.01 and 0.001 and compared it with the dynamic value used in the above experiment with $\beta = 0.008$. Reconstructed images from 100 iterations are shown in figure 9 and the measured RMSE and contrast are shown in figure 10. It is observed that when β was set to a relatively large value, small size and low-contrast inserts were almost lost. On the other hand,

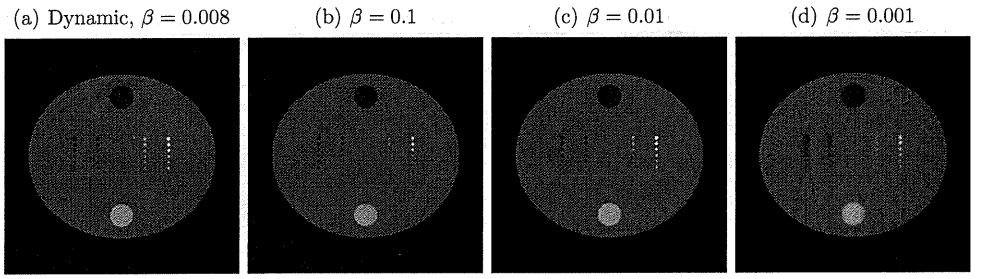

Figure 9. Image reconstructed using OS-iMAP algorithm from 20 projections with (a) dynamic parameter computed using (23) with $\beta = 0.008$ and with static values of (b) $\beta = 0.1$, (c) $\beta = 0.01$ and (d) $\beta = 0.001$.

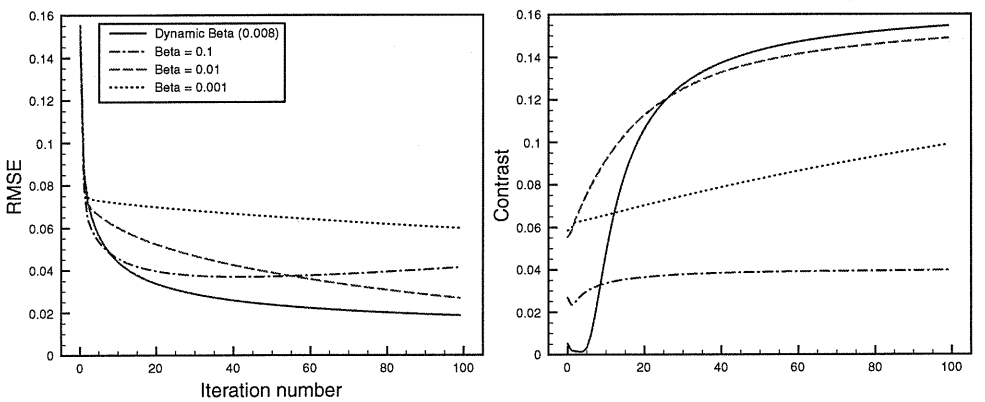

Figure 10. RMSE and contrast computed from images shown in figure 9.

when β is relatively small, streak artifacts were not effectively removed. Moreover, even with intermediate optimized value of $\beta = 0.01$, the image quality measurements are behind the case where the under-relaxation parameter is used as shown in figure 10.

4.3. Real data (head CT)

This experiment was performed using real head CT data obtained from a third-generation scanner. The sampling of the acquired fan-beam data consists of 512(bins) $\times$ 600(views) over 360° which rebinned into a parallel-beam data of 512(bins) $\times$ 600(views) and then manually down-sampled into 75 projection views over 180°. As the blank scan measurements are unavailable, it is estimated from the raw data measurements and a uniform value of 10^4 was used. This study aims to qualitatively compare the OS-iMAP method with the ART-TV method and the conventional FBP. We have limited the number of iterations in both iterative methods to 50 iterations. For the OS-iMAP method, we used $L = 3$, $\beta = 0.02$, $z = (0.0, 0.94, 1.044)\ \text{cm}^{-1}$ and $\omega = (0.02, 0.02, 0.02)$ with five subsets. Reconstructed images are shown in figure 11. The proposed method can significantly reduce streak artifacts using a simple and easy-to-compute intensity prior, which were estimated as the histogram peaks of the FBP image shown in figure 11(d). Image reconstructed using the ART-TV method yields almost no streak artifacts.

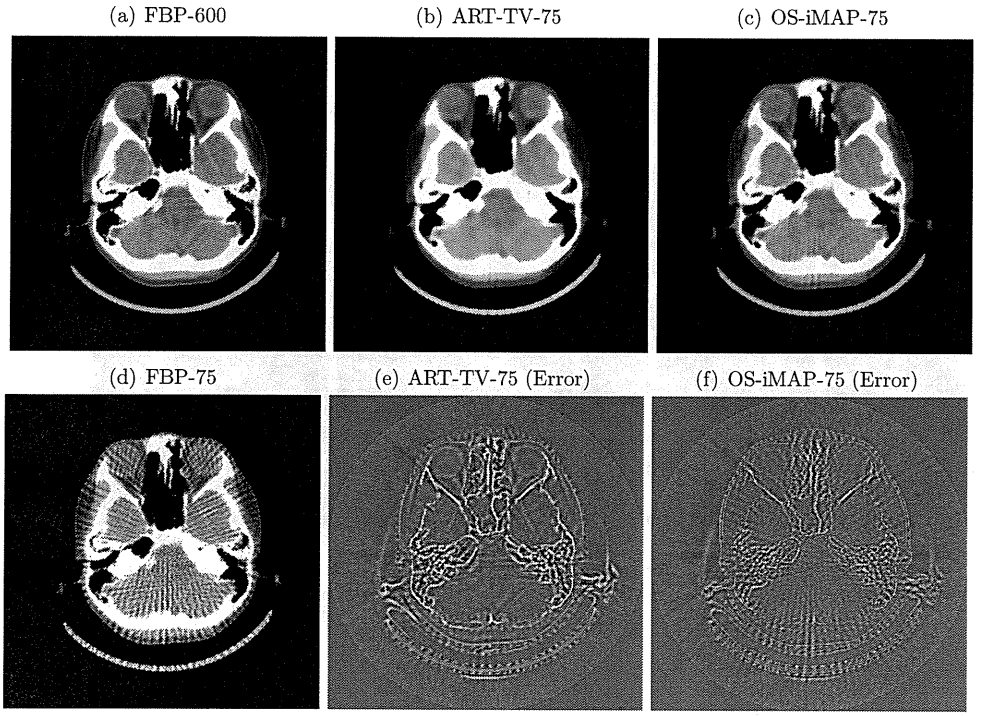

Figure 11. (a) FBP image reconstructed from complete data and images reconstructed from 75 projections using (b) ART-TV, (c) OS-iMAP and (d) FBP algorithms. Error images in (e) and (f) are computed as the difference between (a), (b) and (a), (c), respectively. Display gray scale is (0.85, 1.15) cm^{-1} except for (e) and (f), it is $(-0.1, 0.1)$ cm^{-1}. Displayed images are trimmed to 400×400 pixels for display magnification.

However, it is clear from the error image shown in figure 11(e) that a large portion of image contrast in some regions, especially bone, was lost. On the other hand, the image reconstructed from the OS-iMAP method contains weak artifacts, but the image contrast is highly preserved as shown by the error image in figure 11(f).

4.4. Pseudo-real data (chest CT)

In the pseudo-real data experiment, we used previously reconstructed CT volume to evaluate the proposed method. The original data were scanned using x-ray power of 120 KeV and reconstructed in a volume grid of $512 \times 512 \times 77$ pixels with 0.63 mm pixel size and a slice size of 1.0 mm. A single slice was re-projected into 100 projection views over 180° and reconstruction was implemented using FBP, ART-TV and OS-iMAP methods. The histogram of the FBP image was used to estimate the intensity prior to be used in the OS-iMAP method. We selected four histogram peaks that correspond to the average intensity values of air, lung, soft-tissue and blood. The experiment settings for the OS-iMAP method were $L = 4$, $\beta = 0.02$, $z = (0.0, 0.104, 1.015, 1.05)$ cm^{-1} and $\omega = (0.04, 0.1, 0.03, 0.02)$ with five subsets. A uniform value of 10^5 was used as the blank scan measurement. The reconstruction was implemented using 50 iterations and the results are shown in figure 12. It is observed that the FBP image suffers from severe streak artifacts, which are significantly reduced in

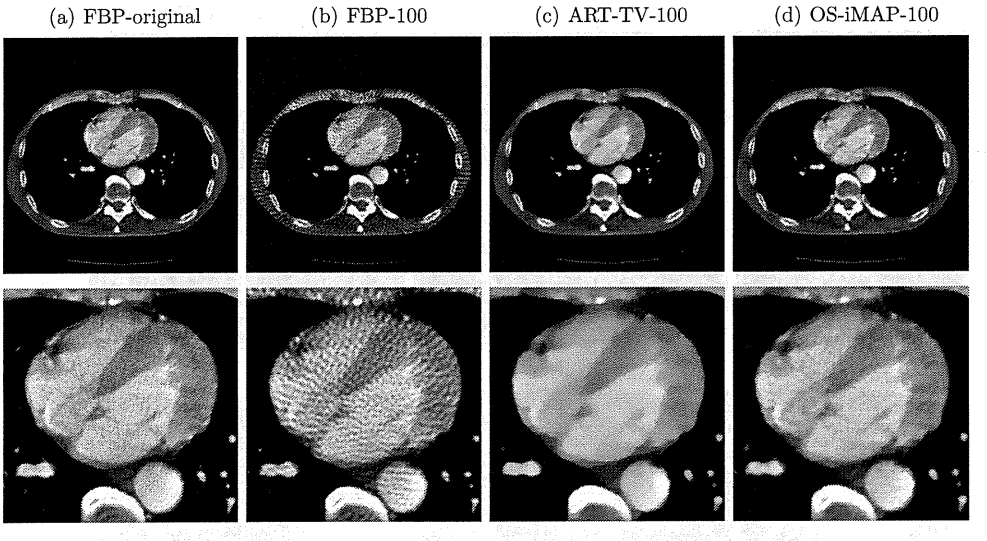

Figure 12. Image reconstruction of the chest CT data. (a) Original slice image and reconstructions from 100 projections using (b) FBP, (c) ART-TV and (d) OS-iMAP methods. A magnification of the cardiac region is shown in the bottom row. Display gray scale is (0.85, 1.3) cm^{-1}.

the ART-TV and OS-iMAP images. However, the well-known carton-like effect of the TV-minimization is avoided in the image reconstructed using the proposed method. Although the OS-iMAP image still suffers from weak artifacts especially in peripheral regions, we expect that these artifacts can be effectively removed with further iterations.

4.5. Real data (high-resolution CT)

In this experiment, we have used a three-dimensional high-resolution CT data obtained from the SPring-8 BL20XU beam line. SPring-8 is a third-generation synchrotron radiation source located in Hyogo, Japan (http://www.spring8.or.jp). The imaged object is a microsample of aluminum alloy (Al–5wt%Cu). In this application, it is desirable to obtain high-quality images from a small number of projection views as the acquired projection data are huge and 3D reconstruction is computationally expensive. The sample of 1 mm diameter was fully scanned with x-ray energy of 35 KeV. The projection data were obtained using a charge-coupled device (CCD) camera of 4000×2624 bins with 2×2 binning mode and 0.5×0.5 μm pixel size. The original parallel-beam projection data used in this experiment were acquired over uniformly spaced 1500 view angles over 180° with an exposure time of 300 ms/view. Two blank scans were obtained through pre- and post-data measurements. The detector array was resampled into 500×328 bins to reduce the computation cost for image reconstruction. The target of this imaging application is to investigate the fine structures of cracks and air holes inside the metal sample, as well as the distribution of different components of the alloy. A central slice reconstructed from 1500 projection views using the FBP algorithm is shown in figure 13.

The projection data were manually down-sampled to 150, 100 and 50 projection views and reconstruction was implemented using FBP, ART-TV and OS-iMAP methods. We have used 8, 12 and 16 iterations for the OS-iMAP algorithm to reconstruct images from 150, 100 and 50 projections, respectively. Almost the same number of iterations was used to obtain the ART-TV image with highest image quality. We have used average intensity values

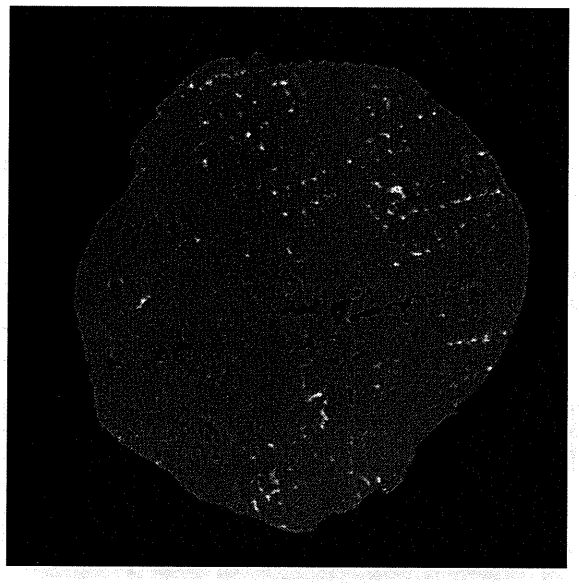

Figure 13. Central slice of the micrometal alloy reconstructed from 1500 projection views using the FBP algorithm.

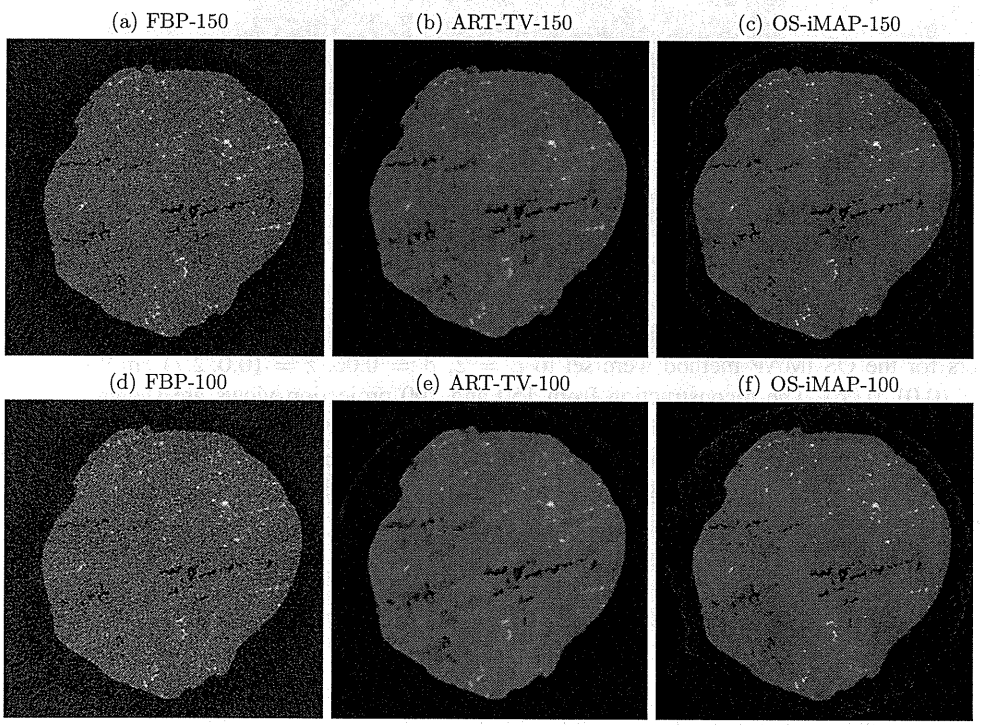

Figure 14. Reconstructed images of the micrometal alloy using FBP (left), ART-TV (middle) and OS-iMAP (right) from 150 (top) and 100 (bottom) projection views.

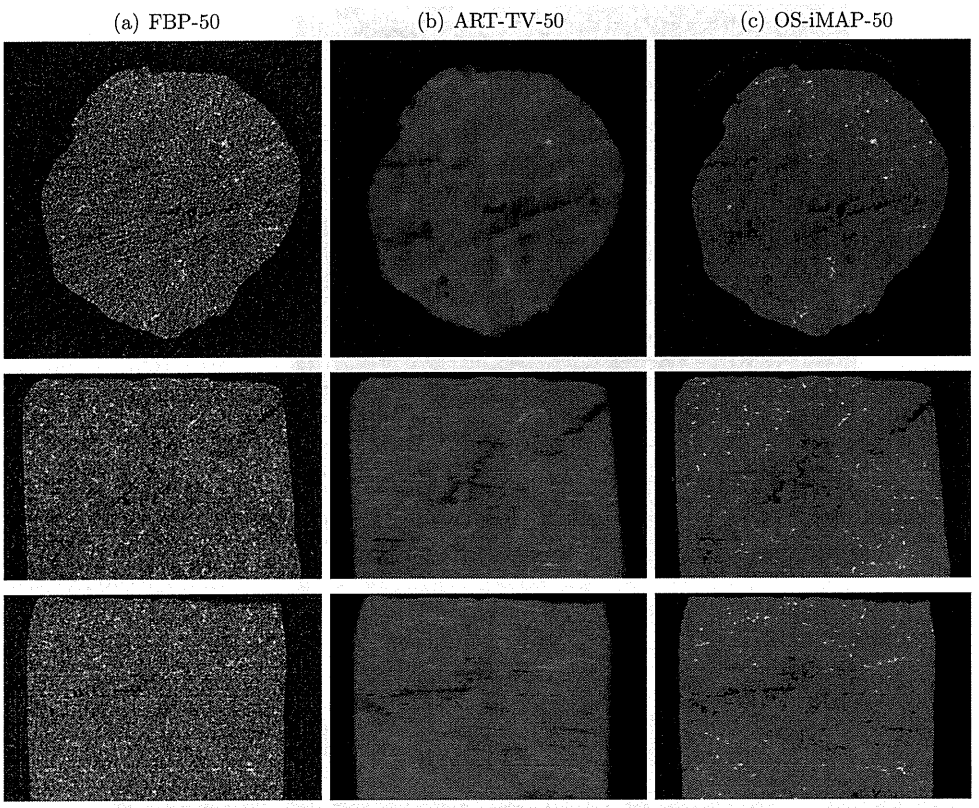

Figure 15. 3D Reconstruction of a micrometal alloy using FBP (left), ART-TV (middle) and OS-iMAP (right) from 50 projection views. Central transverse, coronal and sagittal slices are shown from top to bottom, respectively.

of the homogeneous background (aluminum) and air as the known intensity values. The parameters for the OS-iMAP method were set to $L = 2$, $\beta = 0.06$, $z = (0.0, 2.7)$ cm^{-1} and $\omega = (0.01, 0.06)$. The reconstruction from 150 and 100 projection views are shown in figure 14 and 3D reconstructions of the whole object from 50 projections are shown in figure 15. These results indicate that the image quality is significantly improved by employing the proposed method. Air holes and high contrast metals are clearly observed even by using highly under-sampled projection data.

5. Discussion and conclusion

In this paper, we present a framework for SR from a small number of projection views in x-ray CT. The main contribution is the use of intensity prior as a penalty term in the objective function for image reconstruction. The required intensity information is represented by a small number of intensity values that associated with the homogeneous regions expected in the scanned object. The proposed iMAP algorithm can be easily implemented through a single step of conventional reconstruction followed by a multi-thresholding step in an alternating

manner. Moreover, several possible reconstruction algorithms can be derived using the same approach presented here.

The efficiency of the iMAP algorithm is compared with a recently developed image reconstruction method based on TV-minimization, which is used to solve the problem of image reconstruction from a small number of projections. Both approaches, the iMAP and the ART-TV methods, outperform conventional analytical methods such as FBP in terms of streak artifact reduction. However, small size and/or low contrast structures tend to be lost when the TV minimization is employed, especially when the number of projection views is highly under-sampled, e.g. less than ten projections. It is also known that, in some cases, the TV-minimization may lead to the loss of considerable fraction of the object contrast. Experimental results indicate that the iMAP method effectively preserves small size and/or low contrast image details. Even with real data, where the measured intensity values within the same region is not strictly uniform, the proposed approach outperforms TV minimization in terms of image quality. In addition, the resulting image does not include the carton-like image effect known for the TV-minimization methods.

An important factor in the proposed approach is how to control the hyper-parameter β. During the preliminary experimental studies, we have found that if β is used as a fixed value during iterative reconstruction, the quality of the reconstructed image becomes sensitive to the value of β. For example, if β is selected as a relatively large value, some fine image details would probably be lost. On the other hand, if β is selected as a relatively small value, the effect of multi-thresholding is weakened and the proposed method behaves similar to conventional methods. The use of the under-relaxation approach, such that we start with a relatively large value that gradually decreases with iteration, provides a useful and effective choice. A strong thresholding is required in early iterations to remove streak artifacts and weak thresholding is required later to enforce the data fidelity term to retrieve any possible missing structures.

The proposed method has a large possibility to be used in several low-dose x-ray CT imaging applications. Moreover, the proposed framework is general in such a way that it has a potential to be used in the other data limitation problems. It is interesting to investigate these problems in future work. However, we expect that its validity and power strongly depend on the type of image artifacts to be eliminated using intensity prior.

Acknowledgments

The first author acknowledges support from the Japan Society for the Promotion of Science (JSPS) postdoctoral fellowship (ID no P10052). The authors are very grateful to Dr Hiroyuki Toda (Toyohashi University of Technology, Japan) for providing the high-resolution x-ray CT data, and to the referees for their useful comments.

Appendix. Derivation of equation (21)

For simplicity, let us consider that $L = 1$. Following equations (19) and (20), the minimization of the $\tilde{f}_\beta(\mu; \mu^k)$ in equation (18) is given as

$$q_j = \begin{cases} p_j + 1/(2c_{j,1}) & (p_j < z_1 - 1/(2c_{j,1})) \\ z_1 & (z_1 - 1/(2c_{j,1}) \leqslant p_j \leqslant z_1 + 1/(2c_{j,1})) \\ p_j - 1/(2c_{j,1}) & (p_j > z_1 + 1/(2c_{j,1})) \end{cases} \tag{A.1}$$

$$= \begin{cases} p_j + 1/(2c_{j,1}) (p_j < z_{j,1}^-) \\ z_1 \qquad\qquad (z_{j,1}^- \leqslant p_j \leqslant z_{j,1}^+) \\ p_j - 1/(2c_{j,1}) (p_j > z_{j,1}^+) \end{cases} \tag{A.2}$$

with $z^+_{j,1} = z_1 + 1/(2c_{j,1})$, $z^-_{j,1} = z_1 - 1/(2c_{j,1})$. Similarly, the general multi-thresholding function is given as

$$q_j = \begin{cases} p_j + 1/(2c_{j,1}) & (p_j < z^-_{j,1}) \\ z_1 & (z^-_{j,1} \leqslant p_j \leqslant \min(z^+_{j,1}, s_1)) \\ p_j - 1/(2c_{j,1}) & (z^+_{j,1} < p_j \leqslant s_1) \\ \vdots & \vdots \\ p_j + 1/(2c_{j,l}) & (s_{l-1} < p_j < z^-_{j,l}) \\ z_l & (\max(z^-_{j,l}, s_{l-1}) \leqslant p_j \leqslant \min(z^+_{j,l}, s_l)) \quad (l = 2, \ldots, L-1) \\ p_j - 1/(2c_{j,l}) & (z^+_{j,l} < p_j \leqslant s_l) \\ \vdots & \vdots \\ p_j + 1/(2c_{j,L}) & (s_{L-1} < p_j < z^-_{j,L}) \\ z_L & (\max(z^-_{j,L}, s_{L-1}) \leqslant p_j \leqslant z^+_{j,L}) \\ p_j - 1/(2c_{j,L}) & (z^+_{j,L} < p_j), \end{cases} \tag{A.3}$$

where s_l, $l = (1, \ldots, L-1)$, is the intersection point between the successive ℓ_1 norms associated with intensity values and is given by

$$s_l = (\omega_l z_l + w_{l+1} z_{l+1})/(\omega_l + \omega_{l+1}). \tag{A.4}$$

Therefore, the multi-thresholding function in (A.3) can be written in the following general form:

$$q_j = \begin{cases} p_j + 1/(2c_{j,l}) & (s_{l-1} < p_j < z^-_{j,l}) \\ z_l & (\max(z^-_{j,l}, s_{l-1}) \leqslant p_j \leqslant \min(z^+_{j,l}, s_l)), \quad (l = 1, \ldots, L), \\ p_j - 1/(2c_{j,l}) & (z^+_{j,l} < p_j \leqslant s_l) \end{cases} \tag{A.5}$$

with $s_0 = -\infty$ and $s_L = \infty$.

References

Beekman F J and Kamphuis C 2001 Ordered subsets reconstruction for x-ray CT *Phys. Med. Biol.* **46** 1835–44

Bian J, Siewerdsen J H, Han X, Sidky E Y, Prince J L, Pelizzari C A and Pan X 2010 Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT *Phys. Med. Biol.* **55** 6575–99

Blumensath T and Davies M E 2010 Normalized iterative hard thresholding: guaranteed stability and performance *IEEE J. Sel. Top. Signal Process.* **4** 298–309

Brenner D J and Hall E J 2007 Computed tomography—an increasing source of radiation exposure *N. Engl. J. Med.* **357** 2277–84

Brooks R A, Glover G, Talbert A J, Eisner R L and DiBianca F A 1979 Aliasing: a source of streak in computed tomograms *J. Comput. Assist. Tomogr.* **3** 511–8

Candès E J, Romberg J and Tao T 2006 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52** 489–509

Chen G-H, Tang J and Leng S 2008 Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets *Med. Phys.* **35** 660–3

Daubechies I, Defrise M and De Mol C 2004 An iterative thresholding algorithm for linear inverse problems with a sparsity constraint *Commun. Pure Appl. Math.* **57** 1413–57

Donoho D L 2006 Compressed sensing *IEEE Trans. Inform. Theory* **52** 1289–306

Erdŏgan H and Fessler J A 1999 Monotonic algorithms for transmission tomography *IEEE Trans. Med. Imag.* **18** 801–14

Fahimian B P, Mao Y, Cloetens P and Miao J 2010 Low-dose x-ray phase-contrast and absorption CT using equally sloped tomography *Phys. Med. Biol.* **55** 5383–400

Fessler J A and Hero A O 1995 Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms *IEEE Trans. Image Process.* **4** 1417–29

Galigekere R R, Wiesent K and Holdsworth D W 1999 Techniques to alleviate the effects of view aliasing artifacts in computed tomography *Med. Phys.* **26** 896–904

Hansis E, Schäfer D, Dössel O and Grass M 2008 Evaluation of iterative sparse object reconstruction from few projections for 3D rotational coronary angiography *IEEE Trans. Med. Imaging* **27** 1548–55

Herman G T and Davidi R 2008 Image reconstruction from a small number of projections *Inverse Problems* **24** 045011

Herman G T and Kuba A 1999 *Discrete Tomography: Foundations, Algorithms and Applications* (Boston, MA: Birkhäuser)

Herman G T and Kuba A 2007 *Advances in Discrete Tomography and Its Applications* (Boston, MA: Birkhäuser)

Huesman R H 1977 The effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction *Phys. Med. Biol.* **22** 511–21

Joseph P M and Schulz R A 1980 View sampling requirements in fan beam computed tomography *Med. Phys.* **7** 692–702

Kole J S 2005 Statistical image reconstruction for transmission tomography using relaxed ordered subset algorithms *Phys. Med. Biol.* **50** 1533–45

Kolehmainen V, Siltanen S, Järvenpää S, Kaipio J P, Koistinen P, Lassas M, Pirttilä J and Somersalo E 2003 Statistical inversion for medical x-ray tomography with few radiographs: II. Application to dental radiology *Phys. Med. Biol.* **48** 1465–90

Lange K, Bahn M and Little R 1987 A theoretical study of some maximum likelihood algorithms for emission and transmission tomography *IEEE Trans. Med. Imaging* **6** 106–14

Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **8** 306–16

Lange K and Fessler J A 1995 Globally convergent algorithms for maximum *a posteriori* transmission tomography *IEEE Trans. Image Process.* **4** 1430–8

Leng S, Tang J, Zambelli J, Nett B, Tolakanahalli R and Chen G-H 2008 High temporal resolution and streak-free four-dimensional cone-beam computed tomography *Phys. Med. Biol.* **53** 5653–73

Li M, Kudo H, Hu J and Johnson R H 2004 Improved iterative algorithm for sparse object reconstruction and its performance evaluation with micro-CT data *IEEE Trans. Nucl. Sci.* **51** 659–66

Li M, Yang H and Kudo H 2002 An accurate iterative reconstruction algorithm for sparse objects: application to 3D blood-vessel reconstruction from a limited number of projections *Phys. Med. Biol.* **47** 2599–609

Loris I 2009 On the performance of algorithms for the minimization of ℓ_1-penalized functionals *Inverse Problems* **25** 035008

Mameuda Y and Kudo H 2007 New anatomical-prior-based image reconstruction method for PET/SPECT *Conf. Record of 2007 IEEE Nuclear Science Symp. and Medical Imaging Conf. (Honolulu, USA)* paper no M23-2

Nassi M, Brody W R, Medoff B P and Macovski A 1982 Iterative reconstruction-reprojection: an algorithm for limited data cardiac-computed tomography *IEEE Trans. Biomed. Eng.* **29** 333–41

Natterer F 1986 *The Mathematics of Computerized Tomography* (New York: Wiley)

Ohnesorge B M, Flohr T G, Becker C R, Knez A and Reiser M F 2006 *Multi-Slice and Dual-Source CT in Cardiac Imaging: Principles—Protocols—Indications—Outlook* (Berlin: Springer)

Payot E, Prêteux F J, Trousset Y and Guillemaud R 1997 Generalized support constraint for three-dimensional reconstruction from incomplete Fourier spectra *J. Electron. Imaging* **6** 426–38

Persson M, Bone D and Elmqvist H 2001 Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography *Phys. Med. Biol.* **46** 853–66

Rangayyan R, Dhawan A P and Gordon R 1985 Algorithms for limited-view computed tomography: an annotated bibliography and a challenge *Appl. Opt.* **24** 4000–12

Rashed E A and Kudo H 2011 Row-action image reconstruction algorithm using ℓ_p-norm distance to a reference image *Conf. Record of 2011 IEEE Nuclear Science Symp. and Medical Imaging Conf. (Valencia, Spain)* paper no MIC19-3

Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica* D **60** 259–68

Sauer K, James S Jr and Klifa C 1994 Bayesian estimation of 3D objects from few radiographs *IEEE Trans. Nucl. Sci.* **41** 1780–90

Sidky E Y, Kao C-M and Pan X 2006 Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT *J. x-Ray Sci. Technol.* **14** 119–39

Sidky E Y and Pan X 2008 Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization *Phys. Med. Biol.* **53** 4777–807

Siltanen S, Kolehmainen V, Järvenpää S, Kaipio J P, Koistinen P, Lassas M, Pirttilä J and Somersalo E 2003 Statistical inversion for medical X-ray tomography with few radiographs: I. General theory *Phys. Med. Biol.* **48** 1437–63

Song J, Liu Q H, Johnson G A and Badea C T 2007 Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT *Med. Phys.* **34** 4476–83

Tang J, Nett B E and Chen G-H 2009 Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms *Phys. Med. Biol.* **54** 5781–804

Yu L, Liu X, Leng S, Kofler J M, Ramirez-Giraldo J C, Qu M, Christner J, Fletcher J G and McCollough C H 2009 Radiation dose reduction in computed tomography: techniques and future perspective *Imag. Med.* **1** 65–84

Journal of Cerebral Blood Flow & Metabolism (2012), 1–15

www.jcbfm.com

Review Article

Anesthesia and the quantitative evaluation of neurovascular coupling

Kazuto Masamoto[1,2] and Iwao Kanno[2]

[1]*Center for Frontier Science and Engineering, University of Electro-Communications, Tokyo, Japan;*
[2]*Molecular Imaging Center, National Institute of Radiological Sciences, Chiba, Japan*

Anesthesia has broad actions that include changing neuronal excitability, vascular reactivity, and other baseline physiologies and eventually modifies the neurovascular coupling relationship. Here, we review the effects of anesthesia on the spatial propagation, temporal dynamics, and quantitative relationship between the neural and vascular responses to cortical stimulation. Previous studies have shown that the onset latency of evoked cerebral blood flow (CBF) changes is relatively consistent across anesthesia conditions compared with variations in the time-to-peak. This finding indicates that the mechanism of vasodilation onset is less dependent on anesthesia interference, while vasodilation dynamics are subject to this interference. The quantitative coupling relationship is largely influenced by the type and dosage of anesthesia, including the actions on neural processing, vasoactive signal transmission, and vascular reactivity. The effects of anesthesia on the spatial gap between the neural and vascular response regions are not fully understood and require further attention to elucidate the mechanism of vascular control of CBF supply to the underlying focal and surrounding neural activity. The in-depth understanding of the anesthesia actions on neurovascular elements allows for better decision-making regarding the anesthetics used in specific models for neurovascular experiments and may also help elucidate the signal source issues in hemodynamic-based neuroimaging techniques.
Journal of Cerebral Blood Flow & Metabolism (2012) **0,** 000–000. doi:10.1038/jcbfm.2012.50

Keywords: animal models; awake experiments; neuroimaging

Introduction

Neurovascular coupling consists of three brain cell types: neurons, supporting cells (astrocytes), and vascular cells (vascular smooth muscle, pericyte, and endothelial cells). These cells can be grouped into three conceptual components: neurons, the message senders associated with information processing; supporting cells, the potential transmission sites that mediate vasoactive signals in response to the neuronally derived messages; and vascular cells, the recipients of the signals. After evoked neural activation, vasoactive signals are transmitted directly and indirectly via supporting cells to the vascular cells, which cause redistribution of the local cerebral blood flow (CBF). The hypothetical view of the neurovascular coupling relationship is schematically shown in Figure 1. Great effort has been invested in elucidating the spatiotemporal dynamic functions of neurovascular coupling (Figure 1A), quantitative coupling relationships (Figure 1B), and the mechanisms underlying signal transmission using *in-vivo* animal models (for reviews, see Kleinfeld *et al*, 2011; Attwell *et al*, 2010; Iadecola, 2004; and Lauritzen, 2001).

The anesthesia that has been widely used for studying neurovascular coupling in *in-vivo* animal models has broad action on brain cells that include changes in neural processing, vascular reactivity, and other baseline states (e.g., spontaneous neural activity, cerebral energy metabolism, and baseline CBF). These modulatory effects eventually modify the coupling relationship between neural and vascular responses, and thus, anesthesia is a potential confounder that interferes with the neurovascular coupling relationship. In this article, we review the effects of anesthesia (e.g., anesthetic agents and

Correspondence: Dr K Masamoto, Center for Frontier Science and Engineering, University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan.
E-mail: masamoto@mce.uec.ac.jp
This study was partially supported by Special Coordination Funds for Promoting Science and Technology from the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT).
Received 22 October 2011; revised 8 March 2012; accepted 9 March 2012

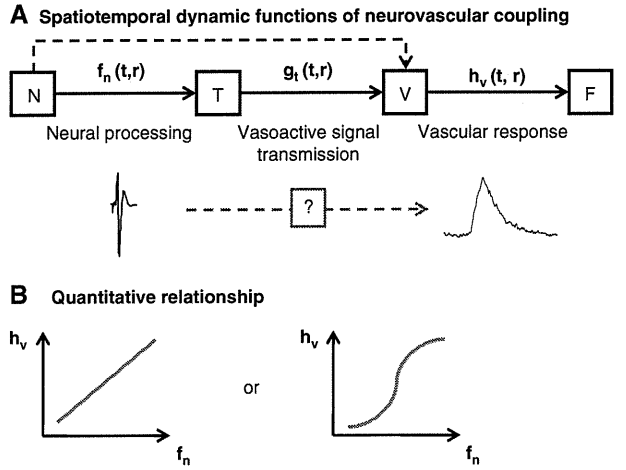

Figure 1 Compartmentalized neurovascular coupling relationship. Neurovascular cells are grouped into three compartments: neurons (N), supporting cells that are potential transmission sites of vasoactive signals (T), and vascular cells (V). The vasoactive signals that are released accompanying neural processing are transferred to the vascular cells directly or indirectly via the supporting cells and help coordinate local cerebral blood flow (CBF) (F). A sequential relationship is depicted (**A**). Each transfer function represents cortical neural processing (f_n), vasoactive signal transmission (g_t), and vascular reactivity (h_v), which are functions of time (t) and space (r). Hemodynamic response, a final output of neurovascular function, is expressed by the convolution of these transfer functions. (**B**) Linear and nonlinear relationships have been observed between neural (f_n) and vascular (h_v) signal changes. However, their relationships to transmission function (g_t) remain uncertain.

depths) on the (1) spatial coordination, (2) temporal dynamics, and (3) quantitative relationships between the neural and vascular responses. To help understand the underlying causes of these anesthesia effects, this review aims to provide comprehensive lists of the actions of anesthesia on (1) general physiology, (2) neurons, (3) supporting cells (vasoactive signal transmission), and (4) vascular cells. This may also aid decision-making regarding anesthetics for specific studies of neurovascular coupling in *in-vivo* animal models.

Why Is Anesthesia Needed?

The need for anesthesia depends on the methodology applied in neurovascular coupling studies (for reviews, see Vanzetta and Grinvald, 2008 and Villringer and Dirnagl, 1995). The electrophysiological techniques that have been used to represent neural activity as a quantitative index to compare vascular responses involves local field potentials (LFPs), and single unit and multiunit spiking activity (Lauritzen, 2001; Logothetis *et al*, 2001). Laminar electrodes have also been used to investigate neurovascular coupling with current source density analysis (Martindale *et al*, 2003). These electrophysiological techniques must fix the electrode(s) at certain locations within the brain, and thus, anesthesia has been used to minimize movement artifacts and eliminate the induced stress to the animals. For anesthesia-free neural recording, telemetry systems and linear arrays of microelectrodes have been introduced in behaving small animals (Schregardus *et al*, 2006) and humans (Keller *et al*, 2009). Electroencephalogram and magnetoencephalogram also provide less-invasive neural measurements and allow concurrent recording with cortical hemodynamics in anesthetized animals (Franceschini *et al*, 2008) and in conscious humans (Rosengarten and Kaps, 2010; Ou *et al*, 2009).

To quantify vascular responses, early studies were conducted with a locally generated hydrogen clearance method with microelectrodes to measure dynamic CBF changes elicited by cortical stimulation (Leniger-Follert and Hossmann, 1979), but this technique is invasive and may disturb cortical microcirculation in the vicinity of the measurement regions. An alternative, less-invasive method, laser-Doppler flowmetry, was introduced to monitor dynamic CBF (Dirnagl *et al*, 1989). This technique has become widely used to examine the temporal and quantitative neurovascular relationship in the anesthetized rodents (Matsuura *et al*, 2000; Matsuura and Kanno, 2001; Ances *et al*, 2000; Ngai *et al*, 1999; Akgoren *et al*, 1996). Some groups have shown that laser-Doppler flowmetry can also be used to monitor longitudinal changes of CBF repeatedly in awake behaving rodents (Takuwa *et al*, 2011; Gu *et al*, 2003). Using either scanning laser-Doppler flowmetry or laser speckle flowmetry, two-dimensional maps of evoked CBF have also been reported in both anesthetized (Du and Pan, 2011; Kannurpatti and Biswal, 2006; Ayata *et al*, 2004; Weber *et al*, 2004) and unanesthetized rodents (Takuwa *et al*, 2011). Optical coherence tomography and functional magnetic resonance imaging (fMRI) techniques provide further layer-specific hemodynamic mapping in anesthetized animal cortex (Chen *et al*, 2009; Jin and Kim, 2008; Maheswari *et al*, 2003; Silva and Koretsky, 2002). These neuroimaging techniques are noninvasive, and measurements can therefore also be acquired in awake conditions with the administration of a light sedative or paralyzing agent (Sicard *et al*, 2003; Peeters *et al*, 2001; Lahti *et al*, 1998). However, caution should be exercised in interpreting the data measured in waking conditions because the obtained signals might be contaminated by restraint-related stress and discomfort or enhanced arousal due to recording noises and motion artifacts (Lahti *et al*, 1998, 1999).

Optical intrinsic signal (OIS) imaging techniques, including near infrared spectroscopy (Fuster *et al*, 2005; Obrig and Villringer, 2003) and diffuse optical imaging (Franceschini *et al*, 2008), also allow for mapping the evoked hemodynamic changes (e.g., blood oxygenation and volume) under both anesthetized (Devor *et al*, 2005; Jones *et al*, 2001) and unanesthetized conditions (Berwick *et al*, 2002;