Few-view CT reconstruction with group-sparsity regularization

Peng Bao | Jiliu Zhou | Yi Zhang

College of Computer Science, Sichuan University, Chengdu 610065, China

Correspondence
Yi Zhang, College of Computer Science, Sichuan University, Chengdu 610065, China.
Email: yzhang@scu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 61671312 and 61302028; National Key R&D Program of China, Grant/Award Number: 2017YFB0802300; Science and Technology Project of Sichuan Province of China, Grant/Award Number: 2018HH0070

Abstract
Classical total variation–based iterative reconstruction algorithm is effective for the reconstruction of piecewise smooth image, but it causes oversmoothing effect for textured regions in the reconstructed image. To address this problem, this work presents a novel computed tomography reconstruction method for the few-view problem called the group-sparsity regularization-based simultaneous algebraic reconstruction technique (SART). Group-based sparse representation, which uses the concept of a group as the basic unit of sparse representation instead of a patch, is introduced as the image domain prior regularization term to eliminate the oversmoothing effect. By grouping the nonlocal patches into different clusters with similarity measured by Euclidean distance, the sparsity and nonlocal similarity in a single image are simultaneously explored. The split Bregman iteration algorithm is applied to obtain the numerical scheme. Experimental results demonstrate that our method both qualitatively and quantitatively outperforms several existing reconstruction methods, including filtered back projection, SART, total variation–based projections onto convex sets, and SART-based dictionary learning.

KEYWORDS
computed tomography, few-view reconstruction, sparse representation, total variation

1 | INTRODUCTION

In recent decades, computed tomography (CT) has been wildly used in clinical diagnosis. However, X-ray radiation may cause cancer and genetic disease. It is hence necessary to reduce the amount of dose during a CT scan. To deal with this problem, many methods have been proposed. These methods can be categorized into 2 groups. The first method is to reduce the operating current, which increases the quantum noise in the projection data. The second method is to decrease the number of sampling views, which generates insufficient projection data, leading to few-view or limited-angle CT. How to reconstruct a high-quality CT image from contaminated or undersampled projection data has attracted a great deal of attention in recent years. In this paper, we focus on few-view CT reconstruction.

Traditional analytic algorithms, such as filtered back projection (FBP), have specific requirements for the completeness of the projection data. Streak artifacts appear when the sampling ratio is low. The iterative reconstruction algorithm is an efficient way to solve this problem. Over the past few decades, the most widely used iterative algorithms for tomography imaging are the algebraic reconstruction technique, simultaneous algebraic reconstruction technique (SART), and expectation maximization. However, when projection views are highly sparse without extra prior information, it is very hard to obtain a satisfactory solution with these classical algorithms. To improve this problem,
additional information is usually merged into the objective function to achieve a robust solution. Compressive sensing (CS) theory has been proved a powerful technique.\textsuperscript{6,7} If an image can be represented sparsely with a certain sparse transform, it can be accurately reconstructed with a probability close to one. Inspired by CS theory, Sidky et al\textsuperscript{8} introduced total variation (TV) minimization into incomplete projection data reconstruction and proposed an efficient iterative reconstruction algorithm based on projection onto convex sets (POCS), called TV-POCS. Although TV-POCS can eliminate streak artifacts to a certain degree, the assumption of TV that the signal is piecewise smooth causes TV-POCS to suffer from oversmoothing effects.\textsuperscript{9} As a result, many variants of TV have been proposed to tackle this problem, such as adaptive-weighted TV,\textsuperscript{10} fractional-order TV,\textsuperscript{11,12} and nonlocal means.\textsuperscript{13,14} Chen and Lang\textsuperscript{15} suggested that a high-quality image can be used to constrain the CS-based reconstruction and this method has been extended to several different reconstruction topics with different prior images. Yu and Wang\textsuperscript{16} constructed the pseudo-inverses of the discrete gradient and discrete difference transforms and adopted a soft threshold-filtering algorithm for few-view CT image reconstruction. Laura proposed an efficient sparse representation (SR) for X-ray medical images, which showed a huge improvement in sparsity than some traditional methods.\textsuperscript{17} An interior SPECT reconstruction method is proposed by inverting a generalized truncated Hilbert transform with singular value decomposition (SVD) in Yu et al.\textsuperscript{18}

Recently, dictionary learning–based methods have been proved effective. In contrast to traditional techniques, which process the image pixel by pixel, a dictionary-based method processes images patch by patch. In 2006, Elad tackled the image denoising problem with a dictionary learning method that uses the K-SVD algorithm.\textsuperscript{19} Mairal et al\textsuperscript{20} extended this method to colour image restoration. For medical imaging problems, the dictionary learning method was first introduced into magnetic resonance imaging. Chen et al\textsuperscript{21} combined the dictionary learning method and TV-based magnetic resonance imaging scheme to further improve image quality. Later, Xu et al.\textsuperscript{22} proposed a low-dose CT image reconstruction method based on dictionary learning. This model introduces the SR constraint of a redundant dictionary as the regularization term, and the performance of a global dictionary and adaptive dictionary was discussed.\textsuperscript{22} Inspired by work combining super-resolution with dual dictionary learning,\textsuperscript{23} Lu et al\textsuperscript{24} respectively used a transitional dictionary for atom matching and a global dictionary for image updating to deal with the few-view problem. Zhao et al\textsuperscript{25} extended this method for spectral CT.

Traditional studies based on dictionary learning have 2 limits. First, the computational burden is very heavy. Second, the relationships among patches are ignored. If the original signals are noisy, the accuracy of sparse coding will decline. Inspired by the research on group sparsity,\textsuperscript{26-29} in this article, we proposed a novel few-view CT reconstruction method based on group-sparsity regularization (GSR) called the GSR-based SART. Instead of processing the image patches sequentially, similar patches are clustered into groups as the basic unit of the proposed group-based SR. Thus, the sparsity and nonlocal similarity in a single image are simultaneously imposed. The remainder of the paper is organized as follows: Section 2 introduces the theory details and numerical scheme for the proposed GSR-SART. Experimental results are provided in Section 3 to demonstrate the performance of our method. The discussion and the conclusion are presented in Section 4.

2 | THE PROPOSED METHOD

2.1 | Imaging model

The general model of CT imaging can be approximately represented as the following discrete linear system:

\[ g = Au, \] (1)

where \( A \) is the system matrix, which is composed of \( M \) row vectors, and \( g \) denotes the measured projection data. Our goal is to reconstruct an image represented by vector \( u \) from the projection data \( g \) and system matrix \( A \). In practice, Equation 1 is known as an ill-posed problem because we consider the insufficient projection data problem caused by few views. This means that we cannot obtain a unique \( u \) by directly inverting Equation 1. To solve the linear system expressed in Equation 1, prior information about the target image is often imposed.

2.2 | SR modeling

Sparse representation for image processing seeks a sparse matrix that contains as few nonzero coefficients as possible to approximately represent the signal. The SR model can be expressed as
\begin{equation}
\{\alpha, D\} = \arg \min_{\alpha, D} |x-D\alpha|^2 + \lambda \|\alpha\|_0,
\end{equation}

where \(x\) denotes an observed signal vector, \(D\) is a dictionary, \(\alpha\) presents the coefficients to represent the signal, \(\lambda\) is a regularization parameter, and \(\|\cdot\|_0\) denotes the \(l_0\) norm. The goal of SR is to seek a sparse vector \(\alpha\) to represent \(x\) for a trained \(D\). To better represent \(x\) with \(\alpha\), it is necessary to choose an effective dictionary \(D\). Some approximation algorithms have been proposed to alternatively optimize \(D\) and \(\alpha\), such as MOD,\textsuperscript{30} K-SVD,\textsuperscript{31} and online learning.\textsuperscript{32}

### 2.3 Group-based SR modeling

Usually, most SR-based methods divide the image into overlapped patches and process them one by one. This operation ignores the nonlocal relationships between different patches. In this paper, we impose the nonlocal similarity constraint into SR to create a GSR-based few-view CT reconstruction method.

First, we divide the CT image \(f\) into \(n\) overlapped patches of size \(\sqrt{P_s} \times \sqrt{P_s}\) using a sliding distance of 4 pixels, where vector \(f_k\) denotes an image patch at location \(k, k = 1, 2, ..., n\). In Figure 1, \(f_k\) is indicated by a small yellow square. As indicated by the red square in Figure 1, \(m\) most similar image patches to \(f_k\) are collected in the \(L \times L\) search window to form a patch set \(G_{fs}\). For simplicity, the Euclidean distance is served as the similarity measurement. Second, all the patches in \(G_{fs}\) are unfolded into vectors and arranged into a matrix of size \(P_s \times m\), denoted by \(f_{G_{fs}}\), which includes each patch in \(G_{fs}\) as its columns. Matrix \(f_{G_{fs}}\) is treated as a group of similar patches. We can then define

\begin{equation}
f_{G_k} = E_{G_k}(f),
\end{equation}

where \(E_{G_k}\) is an operator that extracts group \(f_{G_k}\) from \(f\). Next, we use \(E_{G_k}^T(f_{G_k})\) to denote placing group \(f_{G_k}\) back into the \(k\)th position of the reconstructed image. Now, we can express the whole image \(f\) by averaging all the groups as follows:

\begin{equation}
f = \frac{\sum_{k=1}^{n} E_{G_k}^T(f_{G_k})}{\sum_{k=1}^{n} E_{G_k}^T(1_{P_s \times m})},
\end{equation}

where operator \(/\) indicates the element-wise division of 2 matrices and \(1_{P_s \times m}\) is an all-ones matrix of the same size as \(f_{G_k}\).

Next, we introduce how to obtain the adaptive dictionary for \(D_{G_k}\) each group \(f_{G_k}\). In this model, \(e\) is used to denote the estimate of \(f\), which can be obtained in each iteration, and \(e_{G_k}\) denotes the corresponding estimate of group \(f_{G_k}\). Once we obtain \(e_{G_k}\), we apply SVD to it as follows:

\begin{equation}
e_{G_k} = U_{G_k} \Sigma_{G_k} V_{G_k}^T = \sum_{i=1}^{c} \beta_{G_k,i} u_{G_k,i} v_{G_k,i}^T,
\end{equation}

where \(c\) is the number of atoms in \(D_{G_k}\), \(\beta_{G_k} = \{\beta_{G_k,1}, \beta_{G_k,2}, ..., \beta_{G_k,c}\}\), \(\Sigma_{G_k} = \text{diag}(\beta_{G_k})\) denotes a diagonal matrix for which all the elements except for the main diagonal are zero, \(u_{G_k,i}\) denotes the column of \(U_{G_k}\), and \(v_{G_k,i}^T\) denotes the column of \(V_{G_k}^T\). For group \(f_{G_k}\), each atom of \(D_{G_k}\) is defined as follows:

**FIGURE 1** Group construction. Each patch vector \(f_k\) is extracted from image \(f\) Here, \(G_{fs}\) denotes the set composed of \(m\) most similar patches and \(f_{G_k}\) is a matrix composed of all the patch vectors in \(G_{fs}\).
\[ d_{G_{k,i}} = u_{G_{k,i}} v_{G_{k,i}}^T, \quad i = 1, 2, \ldots, c. \]  

We can then define the expression of the ultimate dictionary \( D_{G_k} \) for group \( f_{G_{k}} \) as follows:

\[ D_{G_k} = \{ d_{G_{k,1}}, d_{G_{k,2}}, \ldots, d_{G_{k,c}} \}. \]

Using \( D_{G_k} \), the GSR model seeks a sparse vector \( \alpha_{G_{k}} = \{ \alpha_{G_{k,1}}, \alpha_{G_{k,2}}, \ldots, \alpha_{G_{k,c}} \} \) to represent \( f_{G_k} \):

\[ f_{G_k} = \sum_{i=1}^{c} \alpha_{G_{k,i}} d_{G_{k,i}}. \]

We can then represent the entire image \( f \) using the set of sparse codes \( \{ \alpha_{G_k} \} \).

\[ f = D_{G} \ast \alpha_{G} = \frac{1}{n} \sum_{k=1}^{n} E_{G_k}^T(f_{G_k}) / \sum_{k=1}^{n} E_{G_k}^T(I_{F_x M}) \]

Here, \( D_{G} \) represents the concatenation of all \( D_{G_k} \), and \( \alpha_{G} \) denotes the concatenation of all \( \alpha_{G_k} \).

\[ 2.4 \quad \text{GSR-SART algorithm} \]

Similar to Sidky and Pan,\(^{33}\) alternating iterations are performed in our algorithm. In the first step, we adopt SART to solve the linear system of Equation 1, which yields a noisy result by minimizing the distance between the measured projection data and the estimated projection data. Specifically, the SART algorithm can be described as follows:

\[ u_{p+1}^j = u_p^j + \omega \sum_{i=1}^{M} A_{i,j} (g_i - \bar{g}_i(u_p)), \]

\[ A_{i,j} = \sum_{j=1}^{N} A_{i,j} \quad \text{for} \quad i = 1, 2, \ldots, M, \]

\[ A_{+,j} = \sum_{i=1}^{M} A_{i,j} \quad \text{for} \quad j = 1, 2, \ldots, N, \]

\[ \bar{g}(u) = Au, \]

where \( A \) is a system matrix of size \( M \times N \) (\( M \) is the total number of projection data, and \( N \) is the total number of image pixels); \( \omega \) is the relaxation parameter; and \( p \) is the iteration number. The second step is to obtain an artifact-reduced result using GSR with the estimated result \( u \) from SART as an initial value. The optimization problem of the second step can be expressed as

\[ \min_{\alpha_{G}} \frac{1}{2} \| f - u \|_2^2 + \lambda \| \alpha_{G} \|_0 \quad \text{s.t.} \quad f = D_{G} \ast \alpha_{G}, \]

where \( \frac{1}{2} \| f - u \|_2^2 \) is an \( l_2 \) data-fidelity term, \( \| \alpha_{G} \|_0 \) is a regularization term, and \( \lambda \) is a regularization parameter. We can obtain adaptive dictionary \( D_{G} \) by applying SVD to the estimate \( e \) of \( f \) according Equations 5, 6, 7. How to calculate \( e \) is given below. Then, Equation 14 becomes

\[ \min_{\alpha_{G}} \frac{1}{2} \| f - u \|_2^2 + \lambda \| \alpha_{G} \|_0 \quad \text{s.t.} \quad f = D_{G} \ast \alpha_{G}. \]

However, Equation 15 is always hard to solve because the \( l_0 \)-norm optimization is nonconvex. In this paper, the split Bregman iteration (SBI) algorithm\(^{34}\) is used to solve this problem. Consider the following constrained optimization problem:
According to SBI, the minimization problem in Equation 16 can be split to subproblems, as shown in Algorithm 1.

**Algorithm 1 Split Bregman Iteration (SBI)**

1. Set \( t = 0 \), choose \( \mu > 0 \), \( b^0 = 0 \), \( x^0 = 0 \), \( y^0 = 0 \)
2. Repeat
3. \( x^{t+1} = \arg \min_x \frac{1}{2} ||x - Gy - b'||_2^2 \)
4. \( y^{t+1} = \arg \min_y \frac{1}{2} ||x^{t+1} - Gy - b'||_2^2 \)
5. \( b^{t+1} = b' - (x^{t+1} - Gy^{t+1}) \)
6. \( t = t + 1 \)
7. Until the stopping criterion is satisfied

We define \( p(f) = \frac{1}{2} ||f - u||_2^2 \), \( q(\alpha_G) = \lambda ||\alpha_G||_0 \). Then, invoking SBI, step 3 of SBI becomes

\[
\begin{align*}
    f^{t+1} &= \arg \min_f \frac{1}{2} ||f - u||_2^2 + \frac{\mu}{2} ||f - DG*\alpha_G^{t'} - b'||_2^2 .
\end{align*}
\]  

(17)

Step 4 of SBI becomes

\[
\begin{align*}
    \alpha_G^{t+1} &= \arg \min_{\alpha_G} \lambda ||\alpha_G||_0 + \frac{\mu}{2} ||f^{t+1} - DG*\alpha_G - b'||_2^2 .
\end{align*}
\]  

(18)

Step 5 of SBI becomes

\[
\begin{align*}
    b^{t+1} &= b' - (f^{t+1} - DG*\alpha_G^{t+1}).
\end{align*}
\]  

(19)

Then, the minimization of Equation 15 is transformed into 2 subproblems concerning \( f \) and \( \alpha_G \). For a given \( \alpha_G \), the \( f \) subproblem in Equation 17 is a strictly quadratic convex optimization problem, which can be defined as

\[
\begin{align*}
    min_f P_1(f^{t+1}) &= min_f \frac{1}{2} ||f - u||_2^2 + \frac{\mu}{2} ||f - DG*\alpha_G^{t'} - b'||_2^2 .
\end{align*}
\]  

(20)

We can obtain a closed solution for Equation 20 by setting the gradient of \( P_1(f^{t+1}) \) to be zero, which can be expressed as

\[
\begin{align*}
    f^{t+1} &= (1 + \mu)E^{-1}(u + \mu(DG*\alpha_G^{t'} + b')) ,
\end{align*}
\]  

(21)

where \( E \) is an identity matrix. For a given \( f \), the \( \alpha_G \) subproblem in Equation 18 can be defined as

\[
\begin{align*}
    min_{\alpha_G} P_2(\alpha_G^{t+1}) &= min_{\alpha_G} \lambda ||\alpha_G||_0 + \frac{1}{2} ||e^{t+1} - DG*\alpha_G||_2^2 ,
\end{align*}
\]  

(22)

where \( e^{t+1} = f^{t+1} - b' \). According to the theorem in Zhang, Zhao and Gao,26 Equation 22 can be transformed into

\[
\begin{align*}
    min_{\alpha_G} P_2(\alpha_G^{t+1}) &= min_{\alpha_G} \frac{\alpha}{\mu} \sum_{k=1}^{n} \left( \frac{1}{2} ||e^{t+1}_{G_k}||_0^2 + \zeta ||\alpha_{G_k}||_0 \right) ,
\end{align*}
\]  

(23)

where \( \zeta = (\lambda \times P \times m \times n) / (\mu \times N) \). Then, Equation 22 is transformed into \( n \) subproblems for all groups \( f_{G_k} \). Because \( f_{G_k} = D_{G_k}\alpha_{G_k} \) and \( e_{G_k} = D_{G_k}\beta_{e_{G_k}} \), the minimization for each group \( f_{G_k} \) can be defined as follows:
\[
\arg \min_{\alpha_{G_k}} \frac{1}{2} \| \alpha_{G_k} \|_2^2 + \beta \| \alpha_{G_k} \|_0 + \xi \| \alpha_{G_k} \|_0.
\] (24)

Algorithm 2 GSR-SART

**Initialization**

Given \( u^0, b^0, f^0, \alpha_G^0, D_G^0, \omega, \lambda, \mu, P, L, m, \xi \)

**Repeat**

**SART step:**

for \( p = 1, 2, ..., P \)

if \( p = 1 \)

\( u^p = \text{SART}(u^0, \omega) \)

else

\( u^p = \text{SART}(u^{p-1}, \omega) \)

end if

end for

if \( u^p < 0 \)

\( u^p = 0 \)

end if

**GSR step:**

for \( t = 0, 1, 2, ..., T \)

update \( f^{t+1} \) by Equation 21

\( e^{t+1} = f^{t+1} - b' \)

for each group \( f_{G_k} \)

update \( D_{G_k} \) by Equations 6, 7

update \( \beta_{G_k} \) by Equation 5

update \( \alpha_{G_k} \) by Equation 25

end for

update \( D_{G_k}^{t+1} \) by concatenating all \( D_{G_k} \)

update \( \alpha_{G_k}^{t+1} \) by concatenating all \( \alpha_{G_k} \)

update \( b^{t+1} \) by Equation 19

end for

\( u^0 = D_G \ast \alpha_G \)

Until the stopping criterion is satisfied

Therefore, a closed solution for Equation 24 can be expressed as follows:

\[
\alpha_{G_k} = \text{hard}(\beta_{G_k}, \sqrt{2 \xi}),
\] (25)

where \( \text{hard}(\cdot) \) denotes the hard thresholding function. Once \( \alpha_{G_k} \) is calculated for all groups, the final solution for the \( \alpha_G \) subproblem is determined.

### 2.5 Summary of the proposed algorithm

Our algorithm is composed of 2 main parts: the SART reconstruction and GSR regularization. We summarize the pseudocode of our GSR-SART algorithm in Algorithm 2.

### 3 EXPERIMENTAL RESULTS

In this section, the results of extensive experiments are reported to validate the proposed method for few-view CT reconstruction. We also determine the impact of the number of best matched patches \( m \) and search window size \( L \).

In the experiments, 3 representative slices, abdominal, pelvic, and thoracic images, are tested to demonstrate the
performance of the proposed GSR-SART. All the images are downloaded from the National Cancer Imaging Archive.† In all the experiments, the image arrays are 20 × 20 cm, and the system projection matrix, in fan-beam geometry, is obtained by Siddon ray-driven algorithm with 64 projection views evenly distributed over 360°. The distance from the source to the rotation center is 40 cm, and the distance from the detector center to the rotation center is 40 cm. We use a flat detector with 512 bins. The images are 256 × 256 pixels. All the experiments are performed in MATLAB 2017a on a PC equipped with an AMD Ryzen 5 1600 CPU at 3.2 GHz and 16 GB RAM.

3.1 | Experimental results

In this subsection, the experimental results of the different clinical images are given. The first 3 experiments are simulated under ideal conditions, which means that the measured projection data are noiseless. The parameters of GSR-SART are set as follows: The patch size is 8 × 8, which means that \( P_s = 64 \); \( m \) is set to 40; and \( L \) is set to 40. For the abdominal images, \( \lambda = 1.5e-5 \) and \( \mu = 0.55 \); for the pelvic image, \( \lambda = 5e-5 \) and \( \mu = 0.1 \); and for the thoracic image, \( \lambda = 1.5e-5 \) and \( \mu = 0.08 \). To further evaluate the proposed GSR-SART method, we compare GSR-SART with FBP, SART, TV-POCS, and dictionary learning-based model (SART-DL). The peak signal-to-noise ratio (PSNR), root-mean-square error (RMSE), and structural similarity (SSIM) are used to quantitatively evaluate the performance of the methods.

The PSNR is defined as

\[
PSNR = 10 \times \log_{10} \left( \frac{(\text{max}(f_i))^2}{\left( \sum_{i=1}^{N} (f_i - f^*_i)^2 \right) / N} \right),
\]

where \( f_i \) is the reconstructed value, \( N \) is the size of \( f \), and \( f^*_i \) is the golden reference value.

The RMSE is defined as

\[
RMSE = \sqrt{\left( \sum_{i=1}^{N} (f_i - f^*_i)^2 \right) / N}.
\]

The SSIM is defined as

\[
SSIM = \frac{2e_f e_r (2\sigma_{fr} + c_2)}{(e_f^2 + e_r^2 + c_1)(\sigma_f^2 + \sigma_r^2 + c_2)},
\]

where \( e_f \) and \( e_r \) are the mean values of \( f \) and \( f^* \); \( \sigma_f \) and \( \sigma_r \) are the standard deviations of \( f \) and \( f^* \), respectively; \( \sigma_{fr} \) is the covariance of them; and \( c_1 \) and \( c_2 \) are constants.

The original abdominal image and reconstruction results are shown in Figure 2. In Figure 2B, the result of FBP contains severe streak artifacts due to the incomplete projection data. It can be observed in Figure 2C that there are still undesirable artifacts in the SART result. In Figure 2D, TV-POCS removes all the streak artifacts, but the reconstructed image of TV-POCS suffers from obvious oversmoothing effects. In Figure 2E, although SART-DL efficiently avoids oversmoothing effects, some fine structures are not well preserved, and some additional artifacts can be noticed. GSR-SART achieves the best visual effect in Figure 2F, which shows that it suppresses most of the artifacts without introducing any other side effects and preserves most of the details.

To further visualize the performance of the methods for the abdominal image, horizontal and vertical profiles of the abdominal image, which are indicated by red lines in Figure 2A, are shown in Figures 3 and 4. The profiles of the original image are given as references. Three regions of interest in each profile are enlarged for better visualization. Several arrows indicate the regions in which discrepancies can easily be identified. Here, the profiles generated by our algorithm are closer to the references. The quantitative evaluations of abdominal image are given in Table 1. It is obvious that our method achieves the best performance for all metrics, which demonstrates the ability of GSR-SART to better reduce artifacts and preserve structure than all the other methods.

†https://imaging.nci.nih.gov/ncia/
In Figure 5, the reconstruction results of the pelvic image are given. Because of the extremely low sampling ratio, it is difficult to obtain useful information from the result of FBP in Figure 5B. The SART method can only remove some of the streak artifacts. In Figure 5D, TV-POCS lowers the spatial resolution while eliminating the streak artifacts, and the

**FIGURE 2** Few-view reconstruction results of the abdominal image from 64 noiseless projections over 360°. The display window is [−150 250] HU. A, Original image and results obtained by B, filtered back projection; C, simultaneous algebraic reconstruction technique (SART); D, total variation–projection onto convex sets; E, SART–dictionary learning; and F, group-sparsity regularization–SART

**FIGURE 3** Horizontal profiles (128th row) of the abdominal image reconstructed by different methods from 64 projection views. A, Overall profiles; B, ROI1; C, ROI2; and D, ROI3 of the overall profiles. DL, dictionary learning; GSR, group-sparsity regularization; POCS, projection onto convex sets; ROI, regions of interest; SART, simultaneous algebraic reconstruction technique; TV, total variation

In Figure 5, the reconstruction results of the pelvic image are given. Because of the extremely low sampling ratio, it is difficult to obtain useful information from the result of FBP in Figure 5B. The SART method can only remove some of the streak artifacts. In Figure 5D, TV-POCS lowers the spatial resolution while eliminating the streak artifacts, and the
edges of the tissues are blurred to different degrees. Note that while SART-DL and GSR-SART eliminate most of the artifacts, GSR-SART maintains the edges clearer than other methods in the region indicated by the red arrow.

To further demonstrate the performance of GSR-SART, the absolute difference images relative to the original image are shown in Figure 6. Here, the loss of structural information in Figure 6B is more than other methods, and the result of SART still have artifacts. In Figures 6C-E, the artifacts are well suppressed, and GSR-SART preserves more details, which can be observed in the lower part of the body, as indicated by the red arrow. The quantitative results, which demonstrate the performance of the proposed algorithm for the pelvic image, are shown in Table 2.

The results of the thoracic image are shown in Figure 7. In Figure 7B, the whole image is filled with streak artifacts, and no clinically valuable structures can be recognized. Although SART removes some artifacts, the spatial resolution is not satisfactory, and the blood vessels in the lungs are clearly blurred. TV-POCS recovers the most vessels in Figure 7D, and the spatial resolution is close to the original image. However, artifacts still exist near the bones, as indicated by red arrows. As shown in Figure 7E, SART-DL introduces some undesired artifacts, which may be introduced by (1) noisy dictionary atoms and (2) inaccurate reconstruction by the overlapped image patches. GSR-SART removes all the artifacts without introducing extra artifacts, and the spatial resolution is closest to the ground truth. The red square

### TABLE 1

| Algorithm       | PSNR   | RMSE   | SSIM   |
|-----------------|--------|--------|--------|
| FBP             | 23.49  | 0.06690| 0.51857|
| SART            | 31.05  | 0.02803| 0.84241|
| TV-POCS         | 36.95  | 0.01421| 0.95726|
| SART-DL         | 41.37  | 0.00854| 0.97245|
| GSR-SART        | **47.41**| **0.00426**| **0.98786**|

Abbreviations: DL, dictionary learning; FBP, filtered back projection; GSR, group-sparsity regularization; POCS, projection onto convex sets; PSNR, peak signal-to-noise ratio; RMSE, root-mean-square error; SART, simultaneous algebraic reconstruction technique; SSIM, structural similarity; TV, total variation. The bold text indicates the best scores of measurements among five methods.
region of Figure 7A is enlarged in Figure 8. It is easy to see that the artifacts and noises are severe, and the spine is distorted heavily. In Figure 8D,E, the noises are still obvious, and the structural details are blurred. Compared with other methods, GSR-SART suppresses more artifacts and noises, and the edges of tissue are better maintained. The quantitative results are shown in Table 3. Consistent with the visual effects, GSR-SART has the best scores for all measurements, and the improvements are impressive.
To demonstrate robustness of our method, we add Poisson noise into projection data to simulate the low-dose situation. In this experiment, the blank scan factor $b_0$ is set to $5 \times 10^6$. The results are shown in Figure 9. In Figure 9A, the result of FBP with 360 views is given as the reference. In Figure 9B, the result is clinically meaningless due to the severe artifacts. In Figure 9C, the strike-like artifacts are not well suppressed. We can still observe that noticeable noises exist in the result of TV-POCS in Figure 9D. Comparing Figure 9E and 9F, although most of the noises and artifacts are eliminated by both methods, it can be sensed that the proposed method has clearer structures and obtains better resolution than SART-DL. The quantitative results, which are shown in Table 4, encourage a consistent trend as visual observation.

### 3.2 Parameter selection

In this subsection, the impacts of several important parameters are analysed.
3.2.1 Effect of $\lambda$

To sense the sensitivity of $\lambda$, experiments are performed with various $\lambda$ ranging from $5e^{-6}$ to $2.5e^{-5}$ in steps of $5e^{-6}$ with a fixed $\mu$. The abdominal image is chosen as the test image. The results are shown in Figure 10. It can be observed that the values of PSNR and SSIM rise dramatically with the increasing of $\lambda$ and reach the peak at $1.5e^{-5}$. While $\lambda$ is bigger than $1.5e^{-5}$, PSNR and SSIM decline slowly. As a result, in our experiment, $\lambda$ is set to $1.5e^{-5}$.

3.2.2 Effect of $\mu$

To explore the sensitivity of $\mu$, experiments are performed with various $\mu$ ranging from 0.3 to 0.6 in steps of 0.05 with a fixed $\lambda$. The abdominal image is chosen as the test image. The results are shown in Figure 11. It is concluded that PSNR and SSIM have similar trend with the change of $\mu$. Both metrics increase as $\mu$ rises and reach the peak when $\mu = 0.55$. After that, both metrics decrease. In our experiment, $\mu$ is set to 0.55.

### TABLE 3

Quantitative results obtained by different algorithms for the thoracic image

|          | PSNR  | RMSE  | SSIM  |
|----------|-------|-------|-------|
| FBP      | 21.99 | 0.07949 | 0.34400 |
| SART     | 30.45 | 0.03004 | 0.79577 |
| TV-POCS  | 38.28 | 0.01219 | 0.95866 |
| SART-DL  | 38.60 | 0.01175 | 0.97412 |
| GSR-SART | **43.53** | **0.00666** | **0.98155** |

### FIGURE 8

Zoom-in region of the thoracic image. The display window is $[-150\ 250]$ HU. Enlarged regions of A, the true image; B, filtered back projection; C, simultaneous algebraic reconstruction technique (SART); D, total variation–projection onto convex sets; E, SART–dictionary learning; and F, group-sparsity regularization–SART.

FIGURE 8

(A) (B) (C)

(D) (E) (F)

Abbreviations: DL, dictionary learning; FBP, filtered back projection; GSR, group-sparsity regularization; POCS, projection onto convex sets; PSNR, peak signal-to-noise ratio; RMSE, root-mean-square error; SART, simultaneous algebraic reconstruction technique; SSIM, structural similarity; TV, total variation. The bold text indicates the best scores of measurements among five methods.
FIGURE 9  The results of noise projection data. The display window is [−150 250] HU. A, Image reconstructed by filtered back projection (FBP) with 360 views and results obtained by B, FBP; C, simultaneous algebraic reconstruction technique (SART); D, total variation-projection onto convex sets; E, SART-dictionary learning; and F, group-sparsity regularization-SART.

TABLE 4  Quantitative results obtained by different algorithms for the noisy abdominal image

| Algorithm       | PSNR   | RMSE    | SSIM   |
|-----------------|--------|---------|--------|
| FBP             | 23.36  | 0.06794 | 0.48931|
| SART            | 30.94  | 0.02837 | 0.83787|
| TV-POCS         | 36.24  | 0.01542 | 0.94161|
| SART-DL         | 38.52  | 0.01185 | 0.95565|
| GSR-SART        | **44.51** | **0.00595** | **0.97736** |

Abbreviations: DL, dictionary learning; FBP, filtered back projection; GSR, group-sparsity regularization; POCS, projection onto convex sets; PSNR, peak signal-to-noise ratio; RMSE, root-mean-square error; SART, simultaneous algebraic reconstruction technique; SSIM, structural similarity; TV, total variation. The bold text indicates the best scores of measurements among five methods.

FIGURE 10  Performance with respect to $\lambda$. PSNR, peak signal-to-noise ratio; SSIM, structural similarity.
3.2.3 | Effect of number of best matched patches $m$

To investigate the sensitivity of $m$, experiments are performed with various $m$ ranging from 30 to 80 in steps of 10 with a fixed search window size of $40 \times 40$. The pelvic image is chosen as the test image. The results are shown in Figure 12. It can be observed that the PSNR reaches peak at $m = 40$ and it slowly declines as $m$ increases further. In contrast to PSNR, the values of SSIM decrease monotonously as $m$ increases. One possible reason for this phenomenon is that PSNR focuses more on the reduction of artifacts and noises, which does not always correspond with structure preservation.

3.2.4 | Effect of search window size $L$

To investigate the sensitivity of $L$, experiments are performed for various $L$ ranging from 30 to 90 in steps of 10 with a fixed number of best-matched patches of 40. The pelvic image is selected as the test image. The results are given in Figure 13. Here, the values of PSNR and SSIM behave similarly. The values first increase and reach a peak when $L = 40$. After that, they decrease slowly. Although there is a rebound after 60, $L = 40$ is still the optimal selection.
DISCUSSION AND CONCLUSION

Despite the rapid development of CT imaging techniques, incomplete projection data reconstruction is still a major problem in this field. In this paper, we propose a novel GSR-based SART algorithm for few-view CT reconstruction called GSR-SART. In this algorithm, we use a GSR model as the regularization term to eliminate streak artifacts and preserve structural details. To further explore nonlocal similarity in a target image, a dictionary of the GSR model, which is adaptively generated at each iteration, is learned from groups composed of similar patches. Three representative clinical slices of different parts of the human body are used to validate the performance of the proposed method. In all the results, our method performs better than all the other popular methods qualitatively and quantitatively under the same sampling conditions. Specifically, GSR-SART demonstrates a superior ability to reduce artifacts and preserve details.

In our experiments, the parameters are manually selected. Grid search is a reasonable method for parameter selection in simulations, but it is not practical in actual situations. We also observe that the optimal selection for each image is different, which makes this problem more complicated. A popular machine learning technique is a possible way to adaptively determine the optimal parameter set by learning from an external dataset.

Another problem we note here is the computational cost. Because of the introduction of group sparsity, the computational complexity of GSR-SART is heavier than that of original dictionary learning–based methods. One possible solution to accelerate the computation is to implement a version that uses parallel computing. Distributed computing, computing clusters, and graphics processing units are 3 alternative approaches.

In conclusion, we are very encouraged by the promising performance of GSR with respect to artifact reduction and detail preservation for few-view CT. These results demonstrate the potential of the group-based SR method for medical imaging. In the future, the proposed network framework will be refined and adapted to deal with other topics in CT imaging, such as low-dose CT, metal artifact reduction, and interior CT. Furthermore, coupling our model with deep learning–based methods is also an interesting topic.36,37

ACKNOWLEDGMENTS

We thank Kim Moravec, PhD, from Edanz Group China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript. This work was supported in part by the National Natural Science Foundation of China under grants 61671312 and 61302028, the National Key R&D Program of China under grants 2017YFB0802300, the Science and Technology Project of Sichuan Province of China under grants 2018HH0070.

ORCID

Peng Bao http://orcid.org/0000-0002-8447-8321
Yi Zhang http://orcid.org/0000-0001-7201-2092

REFERENCES

1. Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. New Engl J Med. 2007;357(22):2277-2284.
2. Chang M. Others. a few-view reweighted sparsity hunting (FRESH) method for CT image reconstruction. J X-ray Sci Technol. 2013;21(2):161-176.
3. Gordon R, Bender R, Herman GT. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. J Theor Biol. 1970;29(3):471.
4. Andersen AH, Kak AC. Simultaneous algebraic reconstruction (SART): a superior implementation of the ART algorithm. Ultrasonic Imaging. 1984;6(1):81-94.
5. Dempster AP. Maximum likelihood from incomplete data via the EM algorithm. J Roy Stat Soc B Met. 1977;39(1):1-38.
6. Candes EJ, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE T Inform Theory. 2004;52(2):489-509.
7. Donoho DL. Compressed sensing. IEEE T Inform Theory. 2006;52(4):1289-1306.
8. Sidky EY, Kao CM, Pan X. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. J X-ray Sci Technol. 2006;14(2):119-139.
9. Tian Z. Others. Low dose CT reconstruction via edge-preserving total variation regularization. *Phys Med Biol.* 2011;56(18):5949-5967.

10. Liu Y. Others. Adaptive-weighted total variation minimization for sparse data toward low-dose x-ray computed tomography image reconstruction. *Phys Med Biol.* 2012;57(23):7923-7956.

11. Zhang Y. Others. Few-view image reconstruction with fractional-order total variation. *J Opt Soc Am A.* 2014;31(5):981-995.

12. Zhang Y. Others. Few-view image reconstruction combining total variation and a high-order norm. *Int J Imag Syst Tech.* 2013;23(3):249-255.

13. Kim H. Total-variation others. Non-local (NLTV) minimization combined with reweighted l1-norm for compressed sensing CT reconstruction. *Phys Med Biol.* 2016;61(18):6878-6891.

14. Zhang Y, Xi Y, Yang Q, et al. Spectral CT reconstruction with image sparsity and spectral mean. *IEEE T Comp Imag.* 2016;2(4):510-523.

15. Chen GJ, Leng S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys.* 2008;35(2):660-663.

16. Yu H, Wang G. A soft-threshold filtering approach for reconstruction from a limited number of projections. *Phys Med Biol.* 2010;55(13):3905-3916.

17. Laura RN. Effective sparse representation of x-ray medical images. *Int J Numer Meth Biomed Engng.* 2886;33(12):2017.

18. Yu H, Yang J, Jiang M, Wang G. Interior SPECT-exact and stable ROI reconstruction from uniformly attenuated local projections. *Commun Numer Meth Eng.* 2010;26(5):693-710.

19. Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE T Image Process.* 2006;15(12):3736-3745.

20. Mairal J, Elad M, Sapiro G. Sparse representation for color image restoration. *IEEE T Image Process.* 2008;17(1):53-69.

21. Chen Y, Ye X, Huang F. A novel method and fast algorithm for MR image reconstruction with significantly under-sampled data. *Inverse Probl Imag.* 2017;4(2):223-240.

22. Xu Q, Yu H, Mou X, Zhang L, Hsieh J, Wang G. Low-dose x-ray CT reconstruction via dictionary learning. *IEEE T Med Imaging.* 2012;31(9):1682-1697.

23. Zhang J, Zhang J, Zhao C, Xiong R, Ma S, Zhao D. Image super-resolution via dual-dictionary learning and sparse representation. In: *IEEE Int Symp Circ S.* Vol. 57; 2012:1688-1691.

24. Lu Y, Zhao J, Wang G. Fair-view image reconstruction with dual dictionaries. *Phys Med Biol.* 2012;57(1):173-189.

25. Zhao B, Ding H, Lu Y, Wang G, Zhao J, Molloi S. Dual-dictionary learning-based iterative image reconstruction for spectral computed tomography application. *Phys Med Biol.* 2012;57(24):8217-8229.

26. Zhang J, Zhao D, Gao W. Group-based sparse representation for image restoration. *IEEE T Image Process.* 2014;23(8):3336-3351.

27. Zha Z, Liu X, Zhou Z, et al. Image denoising via group sparsity residual constraint. In: *Int Conf Acoust Speec.; 2016:1787-1791.*

28. Yang Y, Jia J. An image reconstruction algorithm for electrical impedance tomography using adaptive group sparsity constraint. *IEEE TInstrum Meas.* 2017;66(9):2295-2305.

29. Yang Y, Wu H, Jia J. Image reconstruction for electrical impedance tomography using enhanced adaptive group sparsity with total variation. *IEEE Sens J.* 2017;17(17):5589-5598.

30. Engan K, Aase SO, Hakon Husoy J. Method of optimal directions for frame design. In: *Int Conf Acoust Speec.; Vol. 5; 1999:2443-2446.*

31. Aharon Mm, Elad M, Bruckstein A. Rmk-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE T Signal Process.* 2006;54(11):4311-4322.

32. Mairal J, Bach F, Ponce J, Sapiro G. Online dictionary learning for sparse coding. In: *Int Conf Mach Learn, ICML,* 2009, Montreal, Quebec, Canada, June; 2009:689-696.

33. Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys Med Biol.* 2008;53(17):4777-4807.

34. Goldstein T, Osher S. The split bregman method for l1-regularized problems. *SIAM J on Imaging Sci.* 2009;2(2):323-343.

35. Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys.* 1985;12(2):252-255.

36. Chen H, Zhang Y, Chen Y, et al. LEARN: learned experts’ assessment-based reconstruction network for sparse-data CT. *IEEE T Med Imaging.* 2018;99:1-1.

37. Chen H. Others. Low-dose CT with a residual encoder-decoder convolutional neural network (RED-CNN). *IEEE T Med Imaging.* 2017;36(12):2524-2535.

How to cite this article: Bao P, Zhou J, Zhang Y. Few-view CT reconstruction with group-sparsity regularization. *Int J Numer Meth Biomed Engng.* 2018;34:e3101. [https://doi.org/10.1002/cnm.3101]