

PAPER

Temporal feature prior-aided separated reconstruction method for low-dose dynamic myocardial perfusion computed tomography

To cite this article: Zixiang Chen *et al* 2021 *Phys. Med. Biol.* **66** 045012

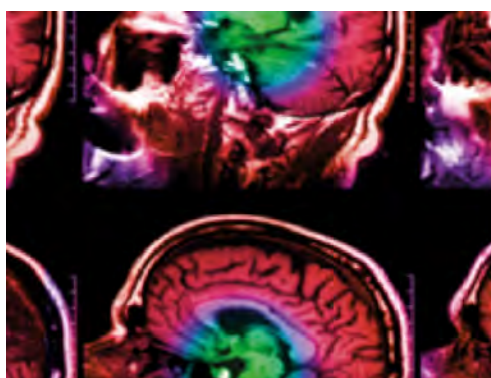
View the [article online](#) for updates and enhancements.

You may also like

- [Oxidative stress and antioxidant defense in response to Dimethyl Phthalate in *Chlorella vulgaris*](#)
Xiaochen Hu, Chenghui Wei, Shuang Wang *et al.*
- [Searching for dark matter particles using Compton scattering](#)
Shang Wang, , Changbo Fu *et al.*
- [Low-dose dynamic myocardial perfusion CT image reconstruction using pre-contrast normal-dose CT scan induced structure tensor total variation regularization](#)
Changfei Gong, Ce Han, Guanghui Gan *et al.*

Recent citations

- [Iterative reconstruction for low-dose cerebral perfusion computed tomography using prior image induced diffusion tensor](#)
Shanzhou Niu *et al*
- [Learning non-local perfusion textures for high-quality computed tomography perfusion imaging](#)
Sui Li *et al*



IPEM | IOP

Series in Physics and Engineering in Medicine and Biology

Your publishing choice in medical physics,
biomedical engineering and related subjects.

Start exploring the collection—download the
first chapter of every title for free.



PAPER

Temporal feature prior-aided separated reconstruction method for low-dose dynamic myocardial perfusion computed tomography

RECEIVED
20 July 2020REVISED
2 December 2020ACCEPTED FOR PUBLICATION
17 December 2020PUBLISHED
2 February 2021Zixiang Chen^{1,2} , Dong Zeng³ , Zhenxing Huang^{1,2}, Jianhua Ma⁴ , Zheng Gu⁵, Yongfeng Yang^{1,2}, Xin Liu^{1,2}, Hairong Zheng^{1,2}, Dong Liang^{1,2}  and Zhanli Hu^{1,2,5} ¹ Lauterbur Research Center for Biomedical Imaging, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, People's Republic of China² Chinese Academy of Sciences Key Laboratory of Health Informatics, Shenzhen 518055, People's Republic of China³ College of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, People's Republic of China⁴ School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, People's Republic of China⁵ Institute of Biomedical Engineering, Shenzhen Bay Laboratory, Shenzhen 518107, People's Republic of ChinaE-mail: zl.hu@siat.ac.cn**Keywords:** dynamic myocardial perfusion computed tomography (DMP-CT), temporal feature, separated reconstruction, time-attenuation curve (TAC), hemodynamic parameter maps**Abstract**

Dynamic myocardial perfusion computed tomography (DMP-CT) is an effective medical imaging technique for coronary artery disease diagnosis and therapy guidance. However, the radiation dose received by the patient during repeated CT scans is a widespread concern of radiologists because of the increased risk of cancer. The sparse few-view CT scanning protocol can be a feasible approach to reduce the radiation dose of DMP-CT imaging; however, an advanced reconstruction algorithm is needed. In this paper, a temporal feature prior-aided separated reconstruction method (TFP-SR) for low-dose DMP-CT images reconstruction from sparse few-view sinograms is proposed. To implement the proposed method, the objective perfusion image is divided into the baseline fraction and the enhancement fraction introduced by the arrival of the contrast agent. The core of the proposed TFP-SR method is the utilization of the temporal evolution information that naturally exists in the DMP-CT image sequence to aid the enhancement image reconstruction from limited data. The temporal feature vector of an image pixel is defined by the intensities of this pixel in the pre-reconstructed enhancement sequence, and the connection between two related features is calculated via a zero-mean Gaussian function. A prior matrix is constructed based on the connections between the extracted temporal features and used in the iterative reconstruction of the enhancement images. To evaluate the proposed method, the conventional filtered back-projection algorithm, the total variation regularized PWLS (PWLS-TV) and the prior image constrained compressed sensing are compared in this paper based on studies on a digital extended cardiac-torso (XCAT) thoracic phantom and a preclinical porcine DMP-CT data set that take image misregistration into account. The experimental results demonstrate that the proposed TFP-SR method has superior performance in sparse DMP-CT images reconstruction in terms of image quality and the analyses of the time attenuation curve and hemodynamic parameters.

1. Introduction

Dynamic myocardial perfusion computed tomography (DMP-CT) is an effective medical imaging approach to quantitatively evaluate functional data on myocardial ischemia and infarction and to diagnose coronary artery disease (CAD), which is a major cause of morbidity and mortality (Ho *et al* 2010, Williams and Newby 2016). The hemodynamic parameter maps calculated from DMP-CT images can offer clinical guidance for medical workers (Uren *et al* 1994, Bamberg *et al* 2011). During the process of DMP-CT image acquisition, two scans are obtained: the baseline scan, taken when the contrast agent has not yet arrived at the cardiac tissue, which

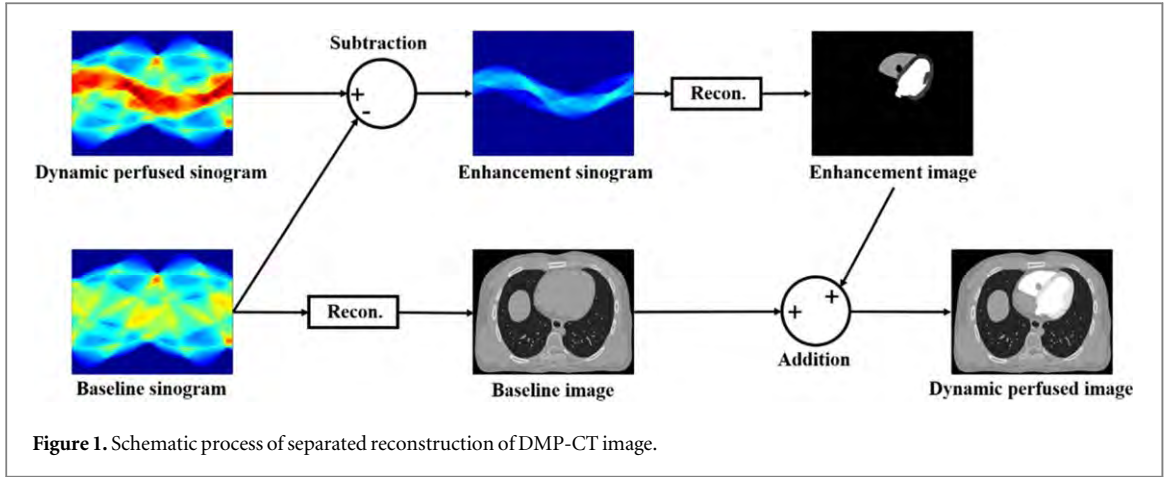
provides the reconstructed baseline image to be subtracted from the dynamic perfusion images for time attenuation curve (TAC) analysis, and the medicine-stressed dynamic perfusion scan performed during the first pass of the contrast agent that provides contrast enhancement (Lee 2002, Becker and Becker 2013), are involved. However, the latent risk of the high-level radiation dose administered by repeated CT scanning during DMP-CT imaging causes great concern regarding the nuclear safety of the examined patients. The average radiation dose for DMP-CT is reported to be approximately 9.15 mSv for the stress scan and approximately 9.09 mSv for the baseline scan (Ho *et al* 2010, Kim *et al* 2013). It is of great importance to reduce the radiation dose from DMP-CT scans in order to promote the use of this powerful diagnostic tool. Sparse acquisition of CT projection data is an effective and feasible method to achieve a low-dose imaging protocol; however, due to data inconsistency, the reconstructed CT images produced by traditional reconstruction algorithms such as the conventional filtered back-projection (FBP) and algebraic iterative reconstruction methods may suffer from severe image errors such as aliasing artifacts caused by data incompleteness (Huang *et al* 2011, Geyer *et al* 2015, Hu *et al* 2019).

Many researchers have focused on advanced algorithms for low-dose DMP-CT image reconstruction. For the low tube current image acquisition protocol, several image reconstruction methods involving innovative regularizations, such as the total variation regularized statistical iterative reconstruction (PWLS-TV), the precontrast normal-dose scan induced structure tensor total variation regularized PWLS method (PWLS-ndiSTV), the simultaneous multiple artifact reduction in tomographic reconstruction method (SMART-RECON) and the motion adaptive sparsity prior involved PWLS method (PWLS-MASP), have been proposed (Tao *et al* 2014, Bian *et al* 2017, Gong *et al* 2017, Li *et al* 2017). For the sparse-view scanning protocol, advanced algorithms have also been investigated and applied. Speidel *et al* proposed a highly constrained back-projection method (HYPR) to reconstruct DMP-CT images from angular staggered sparse projection data using the constraint of an averaged prior image reconstructed via the conventional FBP method from the interleaved assembled sinogram (Speidel *et al* 2007). A classical PWLS-based method for sparse DMP-CT image reconstruction, namely, prior image constrained compressed sensing (PICCS), was proposed by Chen *et al*, in which an interleaved sinogram-reconstructed prior image was used in the process of regularization in the gradient image field (Chen *et al* 2008). Image artifact due to sparse sampling is well suppressed, however, at the cost of image blurring and resolution reduction. Another CS-based method was also applied by Esmaeil Enjilela *et al* for DMP-CT reconstruction, in which the cost function is designed to minimize the summation of the discrete gradients of the objective images, subject to the constraint of image fidelities controlled by the measured sinograms (Donoho 2006, Enjilela *et al* 2018), image reconstructed from 328 views shown comparable quality to normal-dose FBP image. However, we cannot yet promise accurate image reconstruction from ultra-sparse projection data (with tens of projections) via the several previously proposed methods, and a more advanced reconstruction algorithm is needed for clinical application is the task of sparse-view DMP-CT imaging.

Many data-based deep learning methods had been proposed for low-dose CT image reconstruction or post-treatment in recent years, which impules the investigation on low-dose CT imaging technique. Several classical deep learning model, including the directional wavelet convolution neural network (CNN) proposed by Kang *et al* (2017), the modularized deep CNN proposed by Shan *et al* (2019) and the residual encoder-decoder CNN (RED-CNN) proposed by Chen *et al* (2017) were proved to be very effective for restoring high quality images from low-dose CT images that contenting serious noise and artifacts. From another aspect, when considering a sparse-view CT imaging protocol, experimental results attested to that well-designed deep learning network such as the improved GoogLeNet and U-net are capable to reduce artifacts in low-dose CT images (Xie *et al* 2018). The strong capability of deep neural network to extract image features tells the potential values of deep learning methods in low-dose DMP-CT imaging, for both image reconstruction and post-treatment, which should not be neglected.

Prior information is significant for low-dose DMP-CT image reconstruction where serious data inconsistency and incompleteness exist. In the temporal discrete image sequence of DMP-CT, the CT numbers of every tissue of the heart, including the aorta, the ventricles and the myocardium, contain information about the temporal evolution of the contrast agent. This type of information can be used as prior information in the process of image reconstruction. In this study, we proposed a temporal feature prior-aided separated reconstruction method (TFP-SR) for DMP-CT image sequence reconstruction from sparse-view projection data. In our proposed method, the reconstruction of the dynamic perfusion images is separated into two parts: the baseline fraction that represents the preperfused body and the enhancement fraction resulting from the arrival of the contrast agent. The time attenuation evolution information of different tissues of the heart in the DMP-CT image sequence was extracted as the temporal features and innovatively utilized in the process of iterative reconstructions. The experimental results shown that the proposed TFP-SR has the potential to completely replace the commercial used reconstruction method working with normal-dose projection data.

The remainder of this paper is organized as follows: section 2 gives a detailed description of the proposed TFP-SR method and the methods we used for algorithm validation, and the compared methods are listed.



Section 3 and section 4 present the experimental results of our proposed algorithm and the conclusions of our work, respectively. And in section 5, some comments on the proposed TFP-SR method are given.

2. Methods and materials

2.1. Dynamic CT imaging problem formulation

Without loss of generality, the mathematical model of the dynamic CT system can be described by the following linear equation:

$$Y = GX, \quad (1)$$

where $X = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^T\}$ represents the sequence of objects or images containing T temporal frames, and $Y = \{\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^T\}$ is the sequence of sinograms (after system calibration and logarithmic transformation) corresponding to every time frame \mathbf{x}^t . Note that the superscript here denotes the temporal location. G is the system matrix of size $M \times N$ that characterizes the process of forward projection, of which the (i, j) th element represents the contribution coefficient of the j th pixel of image \mathbf{x}^t to the i th pixel of the measured sinogram \mathbf{y}^t . The task of dynamic CT image reconstruction is to estimate the image sequence X from the projection data set Y with the help of knowledge of the imaging system G .

The model for every time frame of the image sequence is similar to that for regular CT imaging, which shows

$$\mathbf{y} = G\mathbf{x}. \quad (2)$$

Therefore, dynamic perfusion CT images can be reconstructed frame by frame with the methods designed for regular static CT.

2.2. Separated reconstruction of DMP-CT images

As is known, after logarithmic transformation of the original detected x-ray signal, every element of the sinogram is the line integral of the attenuation coefficients on the pathway of the corresponding x-ray. Therefore, a linear relationship between the pretreated sinogram and the attenuation coefficients of the scanned object can be stated here: an increase in the CT number of a tissue will lead to a linear increase in the intensity of the corresponding measured sinogram. Furthermore, as discussed in section 1, the x-ray attenuation coefficient of the cardiac tissue where the contrast enhancement agent arrives will increase compared to the same tissue in the baseline state. Then, we obtain the following equalities describing the linear situation in DMP-CT imaging:

$$\begin{cases} \text{Sinogram}_{\text{baseline}} + \text{Sinogram}_{\text{enhancement}} = \text{Sinogram}_{\text{dynamic}} \\ \text{Image}_{\text{baseline}} + \text{Image}_{\text{enhancement}} = \text{Image}_{\text{dynamic}} \end{cases}, \quad (3)$$

where $\text{Sinogram}_{\text{baseline}}$ and $\text{Image}_{\text{baseline}}$ correspond to the baseline fraction of the measured projection data (after system calibration and logarithmic transformation) and the baseline image, respectively, and $\text{Sinogram}_{\text{enhancement}}$ (enhancement sinogram) and $\text{Image}_{\text{enhancement}}$ (enhancement image) correspond to the contrast-enhanced fraction of the measured projection data and the enhancement image, respectively, resulting from perfusion by the contrast agent. $\text{Sinogram}_{\text{dynamic}}$ and $\text{Image}_{\text{dynamic}}$ represent the measured perfusion projection data and the dynamic perfusion image, respectively.

The proposed algorithm in this study is based on the linear relationship described above. For dynamic perfused myocardial CT imaging, the baseline image obtained before dynamic sequential scanning is required for TAC calculation (Lee 2002, Bindschadler *et al* 2014). This baseline image, although containing no

information from the enhanced contrast, shares almost the same structural information with the subsequent dynamic perfused images obtained after the arrival of the contrast agent if there is no dramatic patient motion. Based on this idea and (3), only reconstruction from Sinogram_{enhancement} to Image_{enhancement} is needed if Image_{baseline} is available, and Sinogram_{enhancement} can be obtained by subtracting Sinogram_{baseline} from Sinogram_{dynamic}. In summary, the idea of separated reconstruction of the myocardial perfusion CT images is illustrated in figure 1.

In this study, the resulting perfusion image is composed of the normal-dose baseline image, which is reconstructed by the commercial FBP method with the projection data obtained when the contrast agent has not yet arrived at the cardiac tissue, and the low-dose enhancement image, which is reconstructed by the proposed TFP-SR method with the sparse enhancement sinogram, as demonstrated by (3). To unify the notations for further discussion, the separated reconstruction in this study is described by

$$\text{TFP-SR}_d \mathbf{x} = \text{FBP}_b \mathbf{x} + \text{TFP-SR}_e \mathbf{x}, \quad (4)$$

where the pre-superscript indicates the algorithm via which the corresponding term is reconstructed, and the pre-subscripts 'd', 'b' and 'e' indicate, Image_{dynamic}, Image_{baseline} and Image_{enhancement}, respectively.

2.3. TFP-SR for enhancement image reconstruction

2.3.1. Remodeling of enhancement images

The prior-aided enhancement image reconstruction method is established based on the idea of remodeling the intensities of pixel j of an enhancement image at time t , ${}_e x_j^t$, with a mapping function Γ and the corresponding feature f_j :

$${}_e x_j^t = \Gamma(f_j). \quad (5)$$

The mapping function Γ can be defined by a simple form of weighted sum, which shows

$$\Gamma(f_j) = \mathbf{w}^T \boldsymbol{\theta}(f_j), \quad (6)$$

where \mathbf{w} is a weighting vector, and $\boldsymbol{\theta}$ is an operator that converts the feature f_j to its corresponding transition vector in the transformed data space. The idea of using a transformed feature vector $\boldsymbol{\theta}(f_j)$ instead of f_j is to extend the complexity of the mapping function Γ and guarantee the effectiveness of the remodeling. The weights \mathbf{w} determining how much every element of the transformed feature $\boldsymbol{\theta}(f_j)$ contributes to ${}_e x_j^t$. Based on a consideration that the connections between the current feature and other image features is necessary for the effective design of the remodeling, a novel idea to define the weighting vector \mathbf{w} is to regard it as the weighted sum of all the transformed features $\{\boldsymbol{\theta}(f_1), \boldsymbol{\theta}(f_2), \dots, \boldsymbol{\theta}(f_N)\}$, which gives

$$\mathbf{w} = \sum_{k=1}^N \xi_k^t \boldsymbol{\theta}(f_k), \quad (7)$$

where ξ_k^t is the temporal adaptive weighting coefficient for $\boldsymbol{\theta}(f_k)$. Combining (5)–(7), the pixel intensity ${}_e x_j^t$ can be expressed via the transformed image features by

$${}_e x_j^t = \sum_{k=1}^N \xi_k^t \boldsymbol{\theta}(f_k)^T \boldsymbol{\theta}(f_j). \quad (8)$$

The inner production of two domain transformed variables, $\boldsymbol{\theta}(f_k)$ and $\boldsymbol{\theta}(f_j)$, which are the domain-transformed feature vectors corresponding to enhancement image pixels ${}_e x_k^t$ and ${}_e x_j^t$, respectively, is a typical kernel function of image features f_k and f_j , which is a concept borrowed from machine learning (Hofmann *et al* 2008). By choosing a desired form of the inner product $\boldsymbol{\theta}(f_k)^T \boldsymbol{\theta}(f_j)$, we actually have given a definition of $\boldsymbol{\theta}(f_k)$ and $\boldsymbol{\theta}(f_j)$ implicitly. However, we note that for the application of the proposed TFP-SR method, we need not to figure out what $\boldsymbol{\theta}(f_k)$ and $\boldsymbol{\theta}(f_j)$ are like, instead, the expression of the inner production matters. The distance-based Gaussian function was employed in our work to represent this inner production:

$$\psi(f_j, f_k) = \boldsymbol{\theta}(f_k)^T \boldsymbol{\theta}(f_j) = \exp\left(-\frac{|f_j - f_k|^2}{2\sigma^2}\right). \quad (9)$$

The reason for using the Gaussian function is that by inner product two features f_k and f_j , we are actually building connection between them, and for the effective utilization of the prior information in our TFP-SR method, the Euclidean distance between f_k and f_j is critical and should be considered in the definition of the connection function $\psi(f_j, f_k)$. Subscribing equation (9) into equation (8), we have

$${}_e x_j^t = \sum_{k=1}^N \xi_k^t \psi(f_j, f_k), \quad (10)$$

which can be written in matrix form by

$${}_e\mathbf{x}^t = \Psi \xi^t. \quad (11)$$

where Ψ is the prior matrix, of which the (k, j) th element Ψ_{kj} represents the connection function calculated between f_k and f_j , $\psi(f_j, f_k)$. The remodeling of the enhancement image is achieved through equation (11). And the enhancement image sequence ${}_eX$ can be accordingly expressed by

$${}_eX = \Psi \Xi, \quad (12)$$

where $\Xi = \{\xi^1, \xi^2, \dots, \xi^T\}$ is the sequence of coefficient images.

2.3.2. Prior matrix construction

2.3.2.1. Temporal features

An appropriate definition of the temporal features corresponding to every enhancement image pixel ${}_e\mathbf{x}^t$ is a crucial factor that determines whether the introduction of the temporal information is effective and further the performance of the proposed TFP-SR method. To effectively employ the temporal information naturally contained in the enhancement image sequence, the feature vector f_j of image pixel j is composed of the intensities of the corresponding pixel of all the pre-reconstructed enhancement images in the dynamic sequence ${}_{\text{pre-recon}}{}_eX$, namely

$$f_j = \text{transpose}([{}_{\text{pre-recon}}{}_e\mathbf{x}_j^1, {}_{\text{pre-recon}}{}_e\mathbf{x}_j^2, \dots, {}_{\text{pre-recon}}{}_e\mathbf{x}_j^T]). \quad (13)$$

We note again that the pre-subscript ‘e’ means the enhancement image, the superscript denotes the time frame and the pre-superscript ‘pre-recon’ means that the temporal features were extracted from pre-reconstructed image sequence. In our study, the method used for pre-reconstructing the enhancement images for temporal feature extraction is the total variation regularized maximum likelihood expectation maximization (MLEM-TV) algorithm (Panin *et al* 1999, Chávez-Rivera *et al* 2015).

2.3.2.2. Selection of neighbors

Based on equation (10), the connection function ψ for every two pixels of the target enhancement image should be considered. This will, however, results in a full matrix with size $N^2 \times N^2$ and makes the implementation of the TFP-SR method unfeasible. Efficient implementation requires the prior matrix to be sparse. To achieve this, a k -nearest neighbor principle (k NN) is applied (Friedman *et al* 1977): For a particular feature vector f_j , we chose k feature neighbor vectors (including f_j itself), with the smallest Euclidean distance from f_j , from the temporal feature vectors of all of the pixels to compose the k NN cluster, and the calculation of the elements of the prior matrix was based on

$$\Psi_{j,k} = \begin{cases} \psi(f_j, f_k), & f_k \in k\text{NN cluster of } f_j \\ 0, & \text{otherwise} \end{cases}. \quad (14)$$

And the number of neighbors k involved for prior matrix construction is an important parameter for the proposed TFP-SR method to control the trade-off between the imaging quality and computational cost. Based on the experimental test conducted in our work, $k = 96$ was recommended and used for our experiments.

2.3.3. Prior aided reconstruction

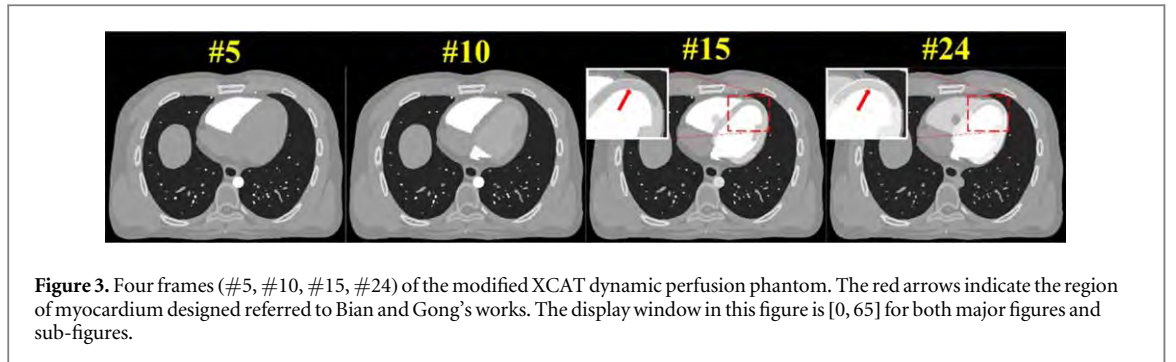
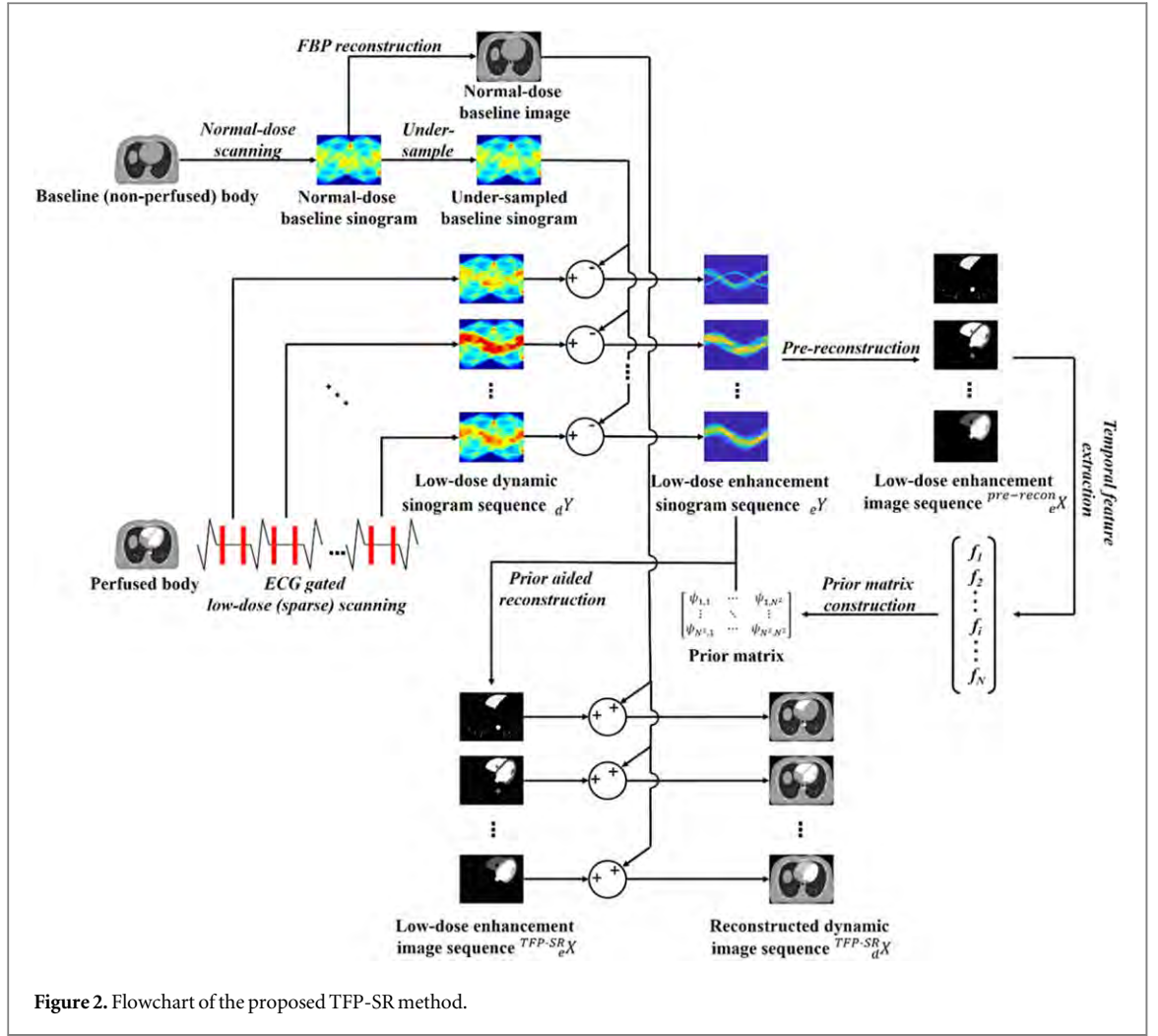
The reconstruction of the enhancement images was based on the MLEM method proposed for transmission and emission computational tomography, of which the updating rule shows (Rockmore and Macovski 1976, Shepp and Vardi 1982, Lange and Carson 1984, Wang *et al* 1998)

$${}_e\mathbf{x}^{t,n+1} = \frac{{}_e\mathbf{x}^{t,n+1}}{G^T \mathbf{1}_M} \cdot \left(G^T \frac{{}_e\mathbf{y}^t}{G {}_e\mathbf{x}^{t,n+1}} \right), \quad (15)$$

where ${}_e\mathbf{x}^t$ is the target image at time frame t with n denoting the iteration step, ${}_e\mathbf{y}^t$ is the corresponding projection data and G is the imaging system matrix as described by equation (2). According to the remodeled image representation given by equation (11), the updating rule of the coefficient image at temporal location t , ξ^t , using the proposed TFP-SR method can be derived from equation (15) by

$$\xi^{t,n+1} = \frac{\xi^{t,n}}{\Psi^T G^T \mathbf{1}_M} \cdot \left(\Psi^T G^T \frac{{}_e\mathbf{y}^t}{G \Psi \xi^{t,n}} \right). \quad (16)$$

And the enhancement image ${}_e\mathbf{x}^t$ is then restored from the reconstructed coefficient image through equation (11).



2.4. Overall process of TFP-SR for DMP-CT image reconstruction

As discussed in section 2.2, a perfused CT image at time t , $d\mathbf{x}^t$, is composed of the baseline image $b\mathbf{x}$ and the corresponding enhancement image $e\mathbf{x}^t$, as shown by (4). In our work, the baseline image was reconstructed by the conventional FBP method from the normal-dose baseline scanning sinogram, while the enhancement images were reconstructed via the proposed TFP-SR method from the low-dose enhancement sinograms. MLEM-TV was employed to pre-reconstruct the enhancement image sequence ${}^{\text{MLEM-TV}}e\mathbf{X}$, from which the feature vectors are extracted in the manner shown in (13). The extracted feature vectors are used to construct the prior matrix Ψ , which is then used in the prior-aided reconstruction iterations to yield the refined enhancement image sequence ${}^{\text{TFP-SR}}e\mathbf{X}$. Finally, the addition of the normal-dose baseline image ${}^{\text{FBP}}b\mathbf{x}$ to every frame in ${}^{\text{TFP-SR}}e\mathbf{X}$ yields the reconstructed dynamic perfused image sequence ${}^{\text{TFP-SR}}d\mathbf{X}$.

In summary, the proposed TFP-SR method can be described by the flowchart and the pseudo code shown in figure 2 and table 1, respectively. We note that the images and sinograms shown in figure 2 are all schematics for

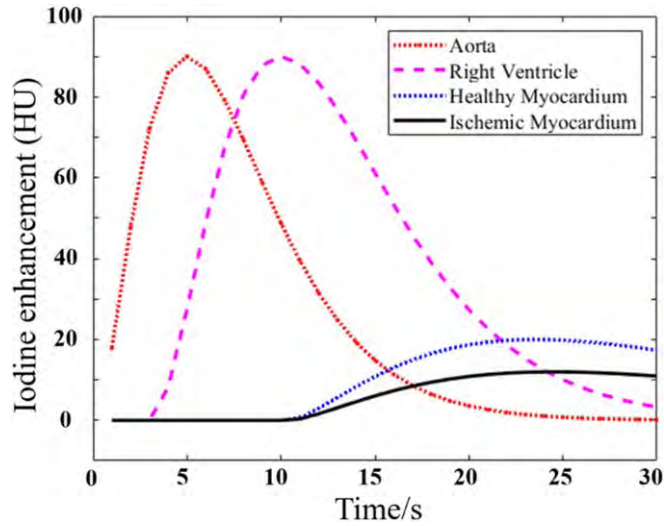


Figure 4. Design TAC for the digital XCAT dynamic phantom.

Table 1. The pseudocode for the TFP-SR method.

Algorithm Temporal feature prior-aided separated reconstruction. (TFP-SR)	
Input: Normal-dose baseline sinogram $^{nd}_b\mathbf{y}$, Low-dose dynamic perfusion sinogram sequence $^{ld}_d\mathbf{Y}$, System matrix of low-dose protocol \mathbf{G}	
Output: Reconstructed dynamic perfusion image sequence $^{TFP-SR}_d\mathbf{X}$	
Parameters: No. of iterations \mathbf{Iter} , TV step size α , No. of neighbors \mathbf{k}	
1. $^{ld}_b\mathbf{y} = \text{Under-sample}(^{nd}_b\mathbf{y});$	//Under-sample the normal-dose sinogram
2. $^{FBP}_b\mathbf{x} = \text{FBP}(^{nd}_b\mathbf{y});$	//Reconstruct the normal-dose baseline image
3. for $t = 1:T$ do	
$^{ld}_e\mathbf{y}^t = ^{ld}_d\mathbf{y}^t - ^{ld}_b\mathbf{y};$	//Sinogram subtraction
endfor	
4. for $t = 1:T$ do	
$^{MLEM-TV}_e\mathbf{x}^t = \text{MLEM-TV}(^{ld}_e\mathbf{y}^t, \mathbf{G}, \alpha, \mathbf{Iter});$	//Pre-reconstruction
endfor	
5. $\Psi = \text{TFP-build}(^{MLEM-TV}_e\mathbf{x}, \mathbf{k});$	//Temporal feature prior construction
6. for $t = 1:T$ do	
$^{TFP-SR}_e\mathbf{x}^t = \text{TFP-SR}(^{ld}_e\mathbf{y}^t, \mathbf{G}, \Psi, \mathbf{Iter});$	//Prior-aided reconstruction
endfor	
7. for $t = 1:T$ do	
$^{TFP-SR}_d\mathbf{x}^t = ^{TFP-SR}_e\mathbf{x}^t + ^{FBP}_b\mathbf{x};$	//Image addition
Endfor	
8. return $^{TFP-SR}_d\mathbf{X} = \{ ^{TFP-SR}_d\mathbf{x}^1, ^{TFP-SR}_d\mathbf{x}^2, \dots, ^{TFP-SR}_d\mathbf{x}^T \}$	

better demonstration of our proposed method. In table 1, the pre-superscript ‘ld’ and ‘nd’ of the sinogram \mathbf{y} represent respectively the low-dose (sparse-view) and the normal-dose cases.

2.5. Data acquisition

2.5.1. Thoracic XCAT phantom study

Figure 3 shows four frames (#5, #10, #15, #24) of the modified thoracic XCAT phantom used for digital phantom study in our work. The time density variations corresponding to different tissues, including the aorta, the right ventricle (RV), the left ventricle (LV), the healthy myocardium (HM) and the modeled ischemic myocardium (IM), are designed based on the physiological driving model to reflect the contrast agent dynamics (Bindschadler *et al* 2014). Thirty temporal frames of dynamic perfused images, corresponding to the end-diastole phase of 30 cardiac cycles, are used to represent an integral process from the first arrival to the total drain of the contrast agent, lasting 30 s (1 frame/s). The designed TACs for the aorta and the healthy and IM are shown in figure 4. The idea of designing of the shape and location of the IM was referred to Bian and Gong’s works (Bian *et al* 2017, Gong *et al* 2017). The TACs are modeled from a gamma variate function, and the simplified formulation developed by Madsen is used (Madsen 1992). The detailed TAC model is given in appendix I. The dynamic perfusion phantom is produced by adding the designed TACs to the baseline phantom. The detailed values of different tissues of the

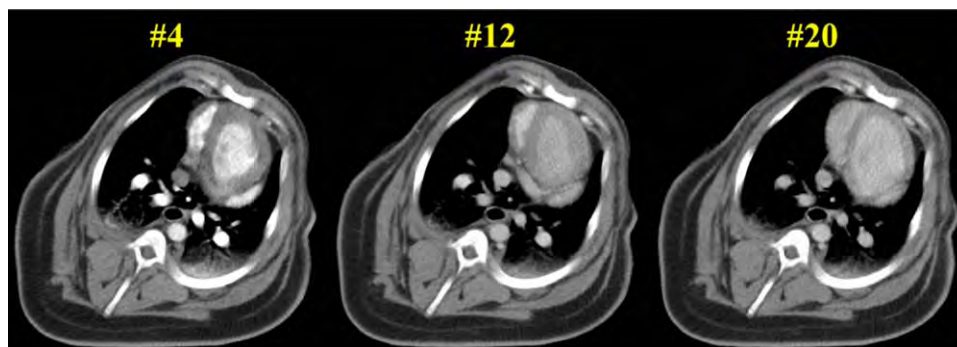


Figure 5. Three frames (#4, #12, #20) of the porcine HDCT image sequence. The display window in this figure is $[-154 \text{ HU}, 256 \text{ HU}]$.

Table 2. Values assigned on different tissues of the baseline phantom.

Tissue	Value
Ventricle	40.32
Lung	9.896
Myocardium	39.98
Muscle	39.92
Rib	55.79
Spinal cord	39.15
Fat	35.3

baseline phantom is given in table 2. The simulative CT projection geometry is designed to represent a typical monoenergetic 2D fan-beam CT scanner that acquires 984 views of projections over 360° with a circular gantry orbit. The source-to-detector distance and the source-to-isocenter distance (SOD) of the simulative scanner are 949 mm and 541 mm, respectively. The fan angle of the x-ray source is 55° , and the detector consists of 888 1 mm channels. For the baseline scan, the normal-dose protocol is applied, and 984 views of projection data are acquired. For dynamic perfused scanning, we apply the low-dose protocol, for which only 41 views of projection data (with a sparse ratio of 24) are measured for every time frame of the dynamic perfused phantoms. The forward projection process is conducted based on an analytical ray-tracing model, and each projection datum along an x-ray through the sectional image is calculated based on the known intensities and intersection areas of the ray with the geometric shape of the object in the sectional image. Similar to that in our previous work (Ma *et al* 2012, Hu *et al* 2016), the noise projections are produced by adding Poisson quantum noise and Gaussian electronic noise to the clean logarithmic projection data. For detector bin i , therefore

$$b_i = \text{Poisson}(I_0 \exp(-\gamma_i)) + \text{Normal}(0, \sigma_e^2), \quad (17)$$

where I_0 is the simulated incident x-ray intensity, namely, the number of photons emitted per ray, and σ_e is the standard deviation of the background electronic noise. The noise sinogram $\tilde{\gamma}_i$ is then calculated as the logarithmic transform of b_i . In our study, I_0 is set to 1.0×10^6 , and σ_e^2 is set to 10.0 for the thoracic phantom study.

2.5.2. Porcine clinical image study

With approval from the Animal Care Committee at the Tianjin Medical University General Hospital (Tianjin China), this study used a sequence of clinically reconstructed dynamic perfused CT images of a mini pig (22.5 kg, female). Before dynamic scanning, ketamine (20 mg kg^{-1}), xylazine (1.5 mg kg^{-1}) and atropine (0.02 mg kg^{-1}) were used for premedical intramuscular administration, and the baseline heart rate and electrocardiograph (ECG) were obtained. An animal aspirator (Matrx VMS Plus VMC, Anesthesia Machines, Midmark Corporation, New York, NY, USA) was used to support mechanical ventilation and anesthesia by sevoflurane (2.5%–3.5% in oxygen). To avoid ventricular fibrillation, intravenous lidocaine injection was conducted. The distinct ST segment elevation shown by the ECG confirmed the acute myocardial infarction. The dynamic scanning was conducted before, during and after the occlusion of the contrast agent, during which the ECG, heart rate, oxygen saturation and arterial blood

pressure were continuously monitored. A 64-slice CT scanner (Discovery CT750 HD, GE Medical Systems) was used under cine mode with a tube current of 100 mA, kVp of 120 kV, gantry rotation time of 0.4 s and reconstruction slice thickness of 5 mm.

The obtained high-dose CT (HDCT) images obtained from the procedure described above are used for the low-dose (sparse) simulation to validate the proposed TFPAR method. Figure 5 shows three frames (#4, #12, #20) of the HDCT porcine clinical image sequence. The simulative projection geometry is designed the same as that described in section 2.3.1, except that 82 views of projection data (with the sparse-sample factor of 12) are acquired from the perfused clinical images.

We note that in the XCAT phantom study, we assume strict image registration between the baseline image and every frame of the dynamic perfusion images. This will facilitate the implementation of the proposed TFP-SR method, but it is not the case in clinical DMP-CT inspection. In the clinical porcine image study, misregistration naturally exists and is considered, which makes the condition closer to the clinical process of DMP-CT imaging.

2.6. Evaluation methods

2.6.1. Evaluation of reconstruction accuracy

To quantitatively evaluate the performance of the proposed TFP-SR method, several widely used indexes are applied. For the digital XCAT phantom study, the peak signal-to-noise ratio (PSNR) and the mean relative root mean square error (rRMSE) are used to evaluate the image reconstruction accuracy (Hore and Ziou 2010). The PSNR is calculated by

$$\text{PSNR}(\mathbf{x}^{\text{recon}}) = 10 \times \log_{10} \left(\frac{\text{MAX}(\mathbf{x}^{\text{true}})^2}{\sum_{j=1}^N (x_j^{\text{recon}} - x_j^{\text{true}})^2 / (N - 1)} \right) \quad (18)$$

and the rRMSE is calculated by Niu *et al* (2014)

$$\text{rRMSE}(\mathbf{x}^{\text{recon}}) = \sqrt{\frac{\sum_{j=1}^N (x_j^{\text{recon}} - x_j^{\text{true}})^2}{\sum_{j=1}^N (x_j^{\text{true}})^2}}, \quad (19)$$

where $\mathbf{x}^{\text{recon}}$ and \mathbf{x}^{true} represent the reconstructed image and the ground truth image, respectively, and N is the total number of pixels involved in the measurement.

2.6.2. Evaluation by myocardial dynamics

The two typical tasks in DMP-CT imaging are to extract the TAC and to calculate the hemodynamic parameter maps, including the myocardial blood flow (MBF), myocardial blood volume (MBV) and mean transition time (MTT). In this study, the TAC of the tissues of interest, namely, the aorta, the HM and the IM, were extracted via regions of interest (ROIs) of size 3×3 , which are marked in figure 7(c). The extracted TAC and the ground truth TAC, as shown in figure 4, are compared to evaluate the accuracy of the reproduced temporal evolution of the contrast agent concentration in the patient's cardiac tissue. The block-circulant single value deconvolution (bSVD) method (Wu *et al* 2003, Bian *et al* 2017) is applied to calculate the hemodynamic parameter maps from the reconstructed dynamic perfusion image sequence, the details of which are supplied in appendix II.

The structural similarity (SSIM) between the calculated MBF, MBV and MTT maps and the ground truth maps, which are calculated from the ground truth dynamic perfusion image sequence, is calculated by Hore and Ziou (2010)

$$\text{SSIM}(M^{\text{recon}}, M^{\text{true}}) = l(M^{\text{recon}}, M^{\text{true}}) \times c(M^{\text{recon}}, M^{\text{true}}) \times s(M^{\text{recon}}, M^{\text{true}}), \quad (20)$$

with $l(M^{\text{recon}}, M^{\text{true}})$, $c(M^{\text{recon}}, M^{\text{true}})$ and $s(M^{\text{recon}}, M^{\text{true}})$ being the luminance index, contrast index and structural information index, respectively, which are calculated by

$$l(M^{\text{recon}}, M^{\text{true}}) = \frac{2\mu_{\text{recon}}\mu_{\text{true}} + C_1}{\mu_{\text{recon}}^2 + \mu_{\text{true}}^2 + C_1}, \quad (21)$$

$$c(M^{\text{recon}}, M^{\text{true}}) = \frac{2\sigma_{\text{recon}}\sigma_{\text{true}} + C_2}{\sigma_{\text{recon}}^2 + \sigma_{\text{true}}^2 + C_2} \quad (22)$$

and

$$s(M^{\text{recon}}, M^{\text{true}}) = \frac{\sigma_{\text{recon-true}} + C_3}{\sigma_{\text{recon}}\sigma_{\text{true}} + C_3}, \quad (23)$$

respectively. In the above three formulas, C_1 , C_2 and C_3 are three small positive constants, μ_{recon} and μ_{true} are the mean values of the calculated map and M^{recon} the ground truth map M^{true} , respectively, and σ_{recon} and σ_{true} are the standard deviations calculated from μ_{recon} and μ_{true} , respectively:

$$\sigma = \frac{1}{N-1} \sum_{j=1}^N (x_j - \mu). \quad (24)$$

$\sigma_{\text{recon-true}}$ is the covariance between M^{recon} and M^{true} that is calculated by

$$\sigma_{\text{recon-true}} = \frac{1}{N-1} \sum_{j=1}^N (M_j^{\text{recon}} - \mu_{\text{recon}})(M_j^{\text{true}} - \mu_{\text{true}}). \quad (25)$$

2.7. Compared algorithms

To evaluate the performance of the proposed TFP-SR method for DMP-CT image reconstruction from sparse projection data, several CT image reconstruction algorithms were compared in our work. These compared methods includes PLWS-based iterative reconstruction algorithms designed specifically for low-dose DMP-CT image reconstruction such as PWLS-TV, PWLS-PICCS and PWLS-ndiSTV (Chen *et al* 2008, Tao *et al* 2014, Gong *et al* 2017) and a deep learning method, namely a typical residual encoder decoder CNN (RED) with 15-layer encoder and 15-layer decoder (Chen *et al* 2017).

2.7.1. PWLS-based methods

The basic target function used for the PWLS-based methods shows

$$\mathbf{x}^* = \underset{\mathbf{x} \geq 0}{\operatorname{argmin}} \{(\mathbf{y} - \mathbf{G}\mathbf{x})^T \Sigma^{-1}(\mathbf{y} - \mathbf{G}\mathbf{x}) + \beta R(\mathbf{x})\}, \quad (26)$$

of which the first term controlling the imaging fidelity and the second term $R(\mathbf{x})$ representing a regularization term and β being a balancing parameter between imaging fidelity and regularization. TV minimization is used as the regularization term in equation (26) for PWLS-TV, which shows

$$R(\mathbf{x}) = \text{TV}(\mathbf{x}). \quad (27)$$

For PICCS and ndiSTV methods, the regularization terms involve the application of prior information about the objective image. $R(\mathbf{x})$ for these two algorithms are written by

$$R(\mathbf{x}) = (1 - \alpha)\text{TV}(\mathbf{x}) + \alpha\text{TV}\mathbf{x}_{\text{nd}} \quad (28)$$

and

$$R(\mathbf{x}) = (1 - \alpha)\text{STV}(\mathbf{x}) + \alpha\text{STV}(\mathbf{x} - \text{nlm}(\mathbf{x}_{\text{nd}})), \quad (29)$$

respectively. α in equations (28) and (29) is a hyper-parameter relatively weighting the normal-dose prior. In equation (29), $\text{STV}(\mathbf{x})$ represents the structural tensor total variation operator of image \mathbf{x} , and $\text{nlm}(\mathbf{x}_{\text{nd}})$ is a non-local mean filtered normal-dose unenhanced image. The method to implement the PWLS-ndiSTV algorithm were described in Gong *et al*'s work (2017).

The hyper-parameter β used for PWLS-TV and PWLS-PICCS methods were set to 100 and 10 for XCAT experiments and porcine clinical image experiments, respectively for the best imaging results, which is equivalently the TV minimization step size during the iterations. For PWLS-ndiSTV, β was set to 0.0604 to be coincide with that used in Gong *et al*'s works. The relative weight α was set to 0.5 for both PICCS and ndiSTV as recommended. We followed the parameters settings given by in Gong *et al*'s work for the implementation of the PWLS-ndiSTV in our work. Similar to that described in the corresponding reference works, an image reconstructed by the FBP algorithm from a set of interleaved sinograms, and the unenhanced image reconstructed via FBP, were used as the prior image for the PICCS and ndiSTV methods, respectively. For all the iterative algorithms, including the three compared methods and the proposed TFP-SR method, 150 iterations were conducted to guarantee the effective convergence.

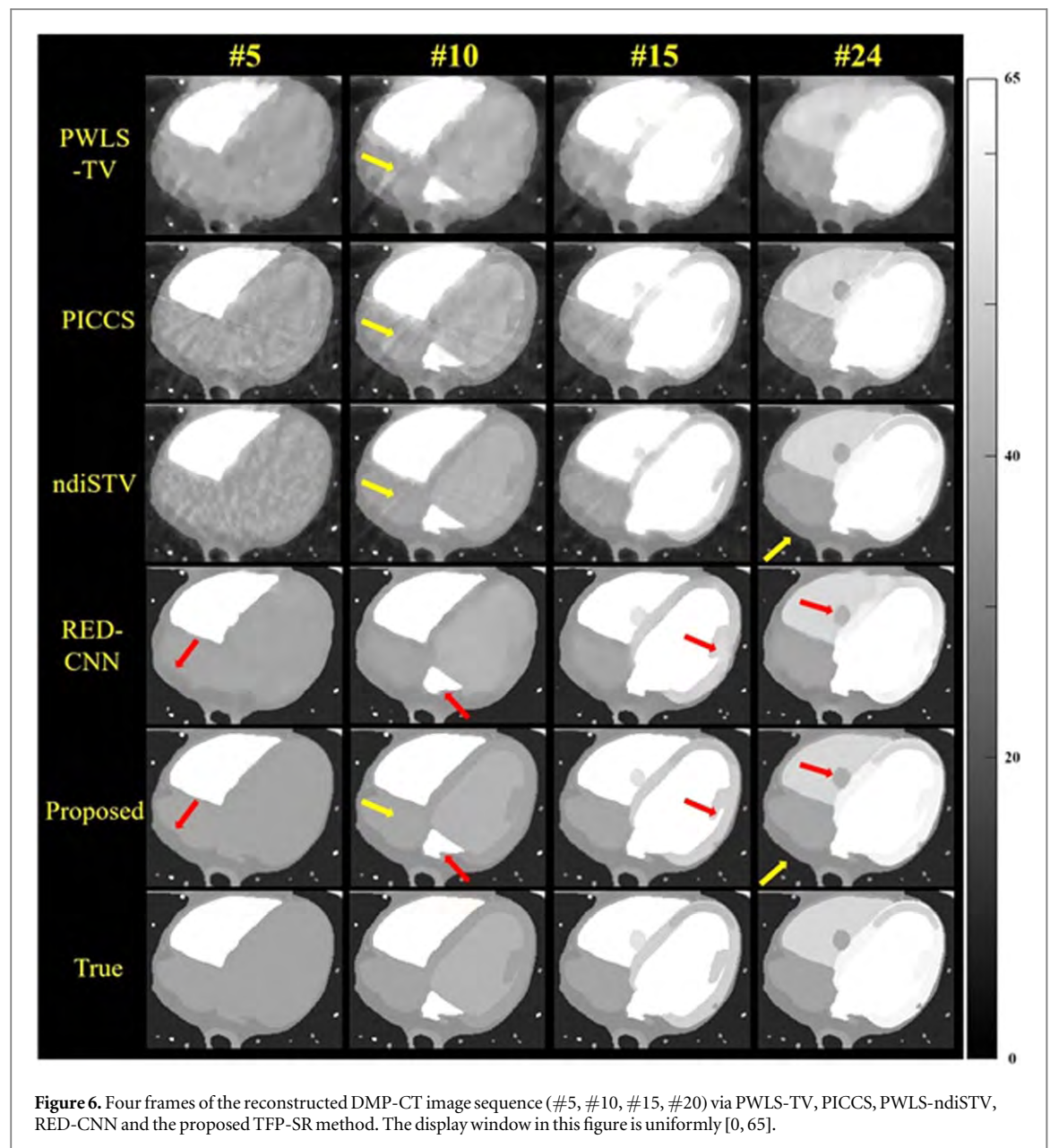


Table 3. Parameters and options of training implementation.

Option or parameter	Setting or value
Padding	0
striding	1
Kernel size	3×3
Patch size	64×64
Loss function	Mean absolute error (MAE)
Optimizer	Adam
Learning rate	Initially 10^{-4} and reduced by half for every 200 epochs

2.7.2. RED-CNN

The RED-CNN involved for algorithms comparison is consisted of 15 encoder layers and 15 decoder layers (Chen *et al* 2017). For the XCAT cardiac phantom study, 12 000 image pairs (one low-dose image reconstructed via FBP and one ground truth phantom count one pair) from three simulative perfused patients were used as the training set. The dynamic perfusion for the training data was designed to be identical to that of our XCAT phantom experiments. For the porcine clinical image study, the size of data set for network training is limited, for which we implemented a data augmentation on the porcine clinical image data set, and finally 6000 image

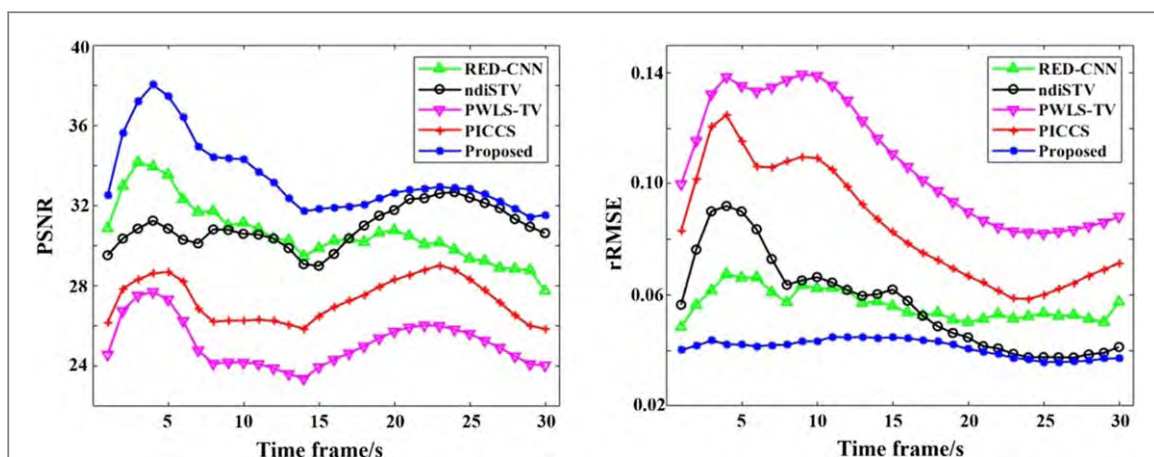


Figure 7. Comparison of PSNR and rRMSE of every frame of the reconstructed DMP-CT image sequences.

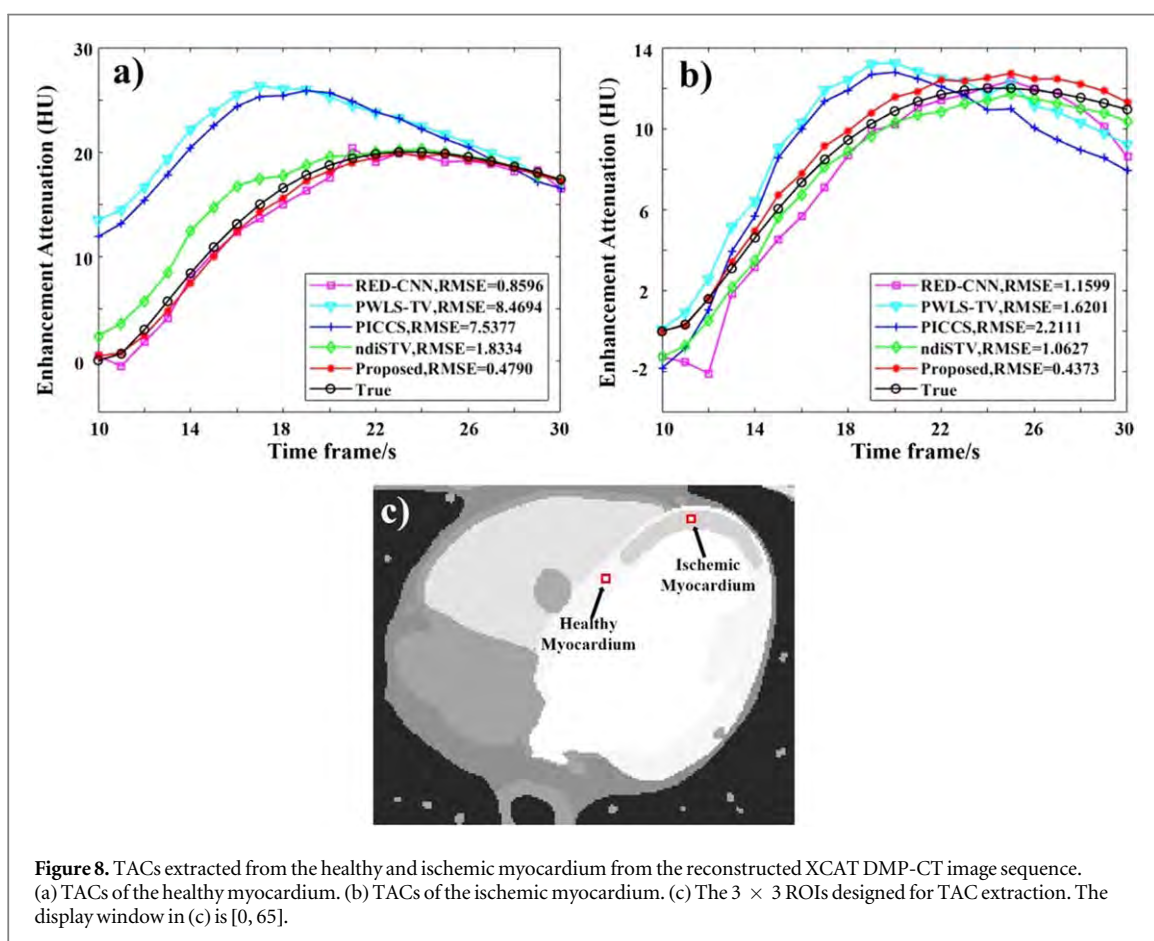


Figure 8. TACs extracted from the healthy and ischemic myocardium from the reconstructed XCAT DMP-CT image sequence. (a) TACs of the healthy myocardium. (b) TACs of the ischemic myocardium. (c) The 3×3 ROIs designed for TAC extraction. The display window in (c) is [0, 65].

pairs (one low-dose image reconstructed via FBP and one high-dose image count one pair) were used as the training set (with the tested 25 images eliminated). The deep learning experiments was implemented in Pytorch on a TITAN 2080 GPU. Other parameters and options associated with the training implementation of the RED-CNN is listed in table 3.

3. Results and discussion

3.1. Thoracic XCAT phantom study

3.1.1. Visual inspection

Figure 6 shows the ROI of the heart of four temporal frames of the DMP-CT image sequence (#4, #10, #15, #24) reconstructed by PWLS-TV, PICCS, PWLS-ndiSTV, RED-CNN and the proposed TFP-SR method.

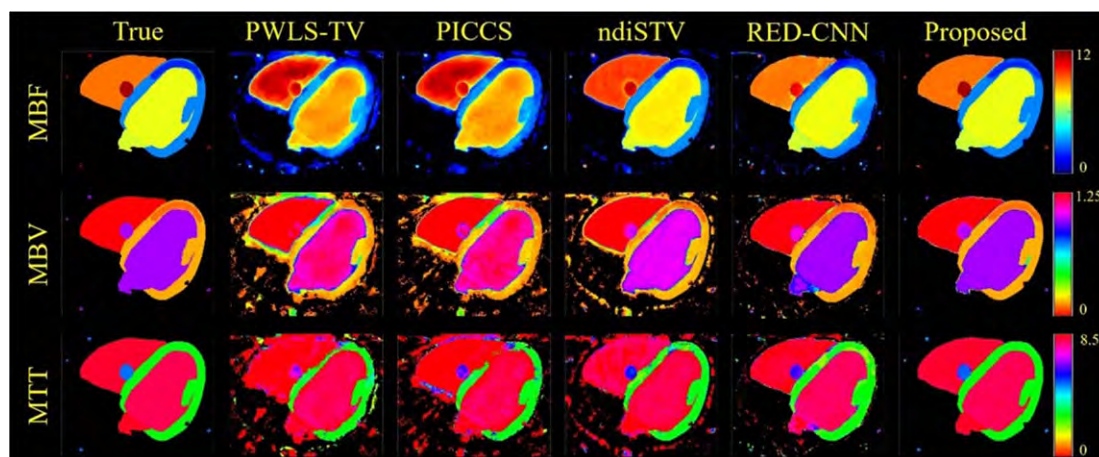


Figure 9. Calculated MBF, MBV and MTT maps based on the DMP-CT image sequences reconstructed by the FBP, PWLS-TV, PICCS and TFP-SR methods.

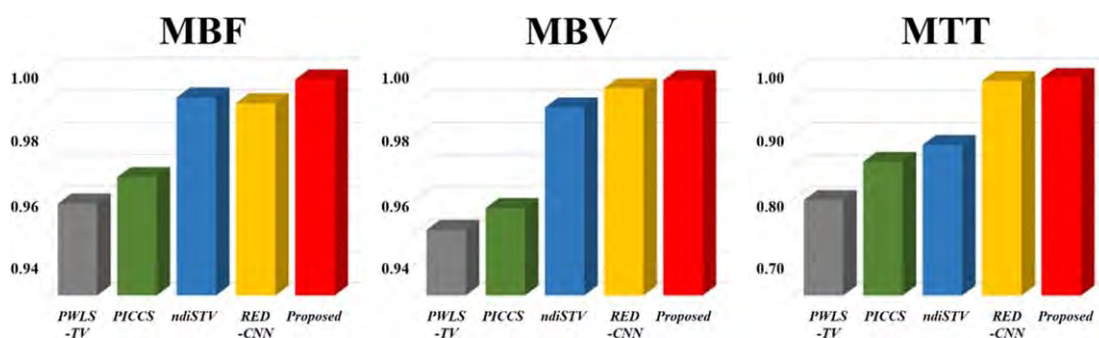


Figure 10. SSIMs of the calculated hemodynamic parameter maps.

Visually speaking, the imaging quality using TFP-SR is remarkably higher than the compared methods. The contrast enhancement in the reconstructed images is restored better, and the boundaries separating different tissues of the heart are preserved better in the images given by TFP-SR than that given by the RED-CNN model (refer to the red arrows), which is a very important factor for clinical diagnosis.. Furthermore, compared with the iterative methods, image artifacts are significantly suppressed (refer to the yellow arrows), indicating the strong capability of the TFP-SR method to solve the problem of sparse-view DMP-CT image reconstruction.

3.1.2. Image reconstruction accuracy

The PSNR and rRMSE values of different temporal frames calculated between the ground truth ROIs and the reconstructed ROIs are given in figure 7. The performance of reconstruction algorithms in the term of imaging accuracy is well illustrated, and the largest PSNR values and the smallest rRMSE values indicates the highest reconstruction accuracy of the proposed TFP-SR method. CT imaging error can significantly affect the further clinical analysis, and the best imaging accuracy of TFP-SR, from this aspect, may strongly states its value in clinical DMP-CT application.

3.1.3. Evaluation by the dynamics of MBF

Figures 8(a) and (b) shows the comparisons between the TACs of respectively the HM and the IM extracted from the DMP-CT image sequences reconstructed by PWLS-TV, PICCS, PWLS-ndiSTV, RED-CNN and the proposed TFP-SR algorithms. The TACs were extracted from two corresponding 3×3 ROIs indicated by figure 8(c). The TACs of both the healthy and ischemic myocardia provided by TFP-SR are the closest to the designed ground truth myocardial TACs, which is coincide with the smallest RMSE values given in the legends of the figures. This tells that the dynamic evolution of the contrast agent in restored with highest accuracy via the TFP-SR method, from which a more reliable clinical hemodynamic analysis can be carried out.

Figure 9 shows the myocardial hemodynamic maps, including the MBF, MBV and MTT images, calculated from the image sequences reconstructed by different algorithms. Compared to other algorithms, the proposed TFP-SR method provides hemodynamic maps with highest quality. We note that the quality of the restored

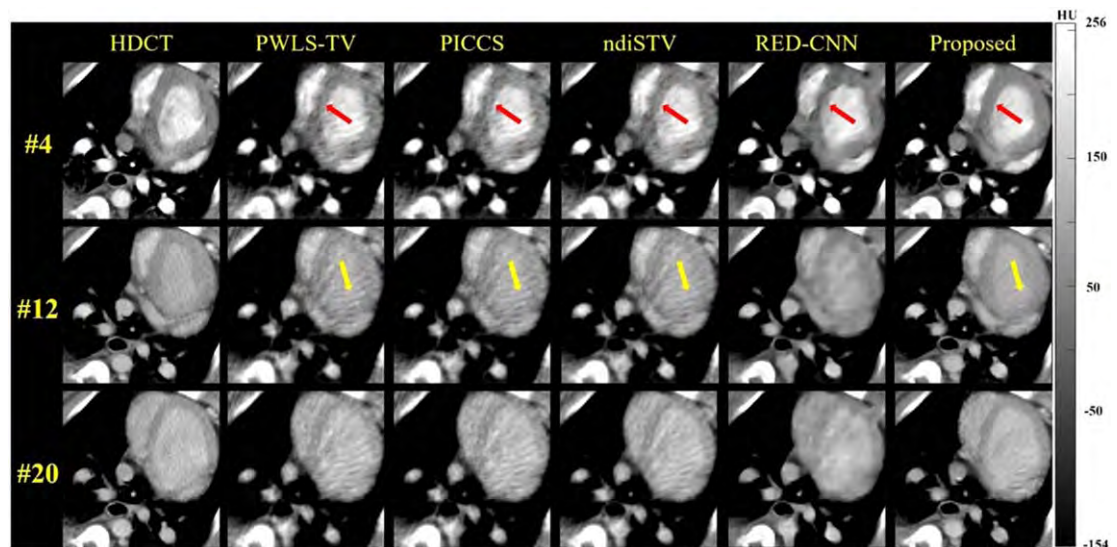


Figure 11. Three frames (#4, #12, #20) of the reconstructed porcine DMP-CT image sequence. The display window is uniformly [−154 HU, 256 HU] for the porcine images.

maps has high correlation with the accuracy of the extracted TACs, which is further related to the quality of image reconstruction. Quantitative evaluation of the calculated hemodynamic maps is given in figure 10. According to the results, PWLS-TV and PICCS are not comparable to other algorithms involved, and the TFP-SR method gives the highest SSIM values for all the MBF, MBV and MTT maps calculated. The hemodynamic maps extracted from DMP-CT image sequence is one of the most direct mediums for clinical diagnosis of CAD and other cardiac blood diseases, therefore, high quality maps are required. From this aspect, TFP-SR can be the most effective algorithm for the task of DMP-CT imaging.

3.2. Porcine clinical images study

3.2.1. Visual inspection

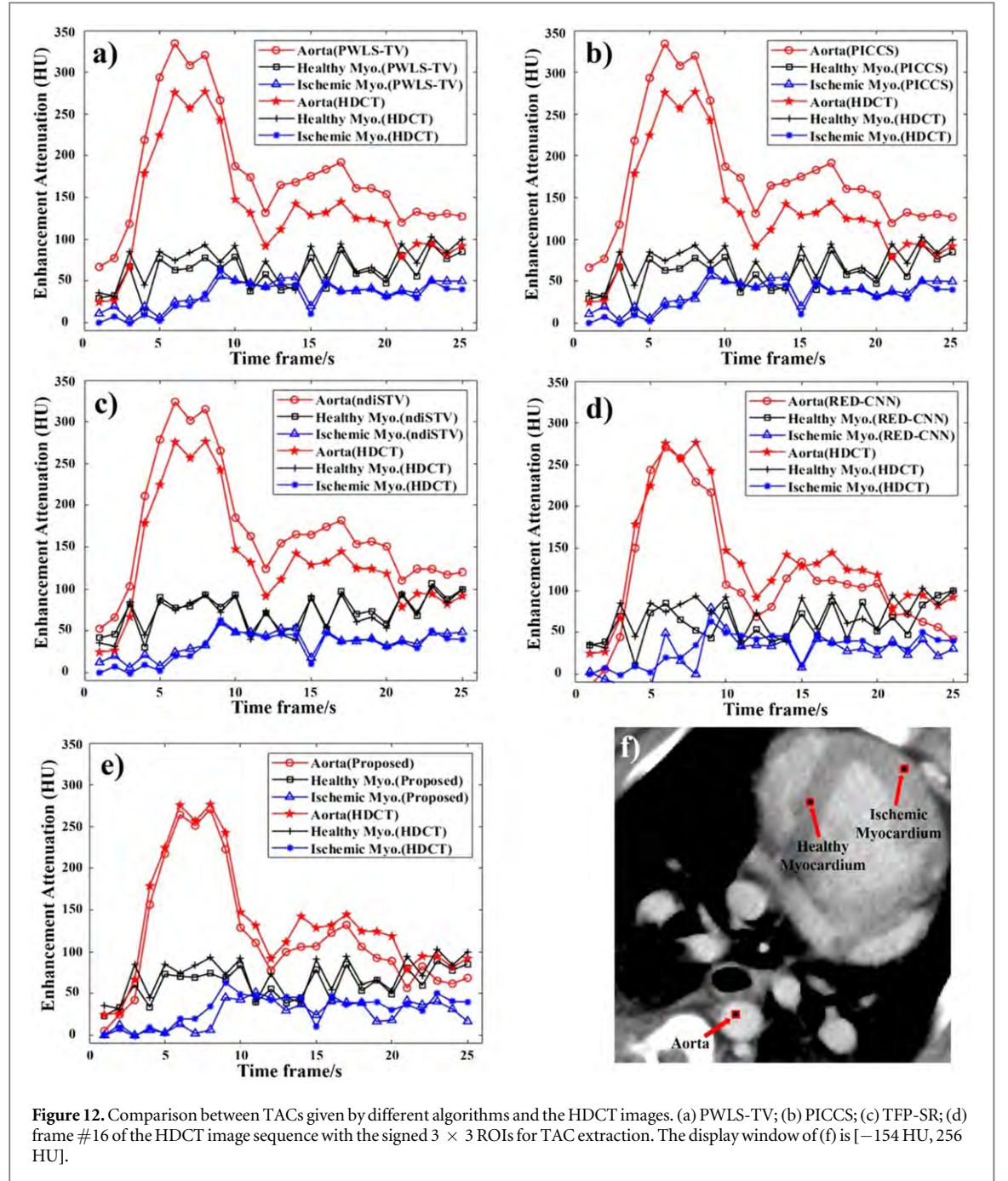
Figure 11 shows three frames (#4, #12, #20) of the reconstructed porcine DMP-CT image sequence reconstructed via PWLS-TV, PICCS, PWLS-ndiSTV, RED-CNN and the proposed TFP-SR methods. Compared to other algorithms, the image artefact are well suppressed by the TFP-SR method, as indicated by the yellow arrows in the second row of figure 11 (#12). Besides, the TFP-SR methods has the best performance in edge preservation, the visual resolution of the myocardial edge shows higher than the compared methods, as indicated by the red arrows in #4. The perfusion-induced contrast enhancement, indicated by the HDCT images, is completely preserved in the TFP-SR reconstructed images.

3.2.2. Evaluation by the dynamics of MBF

Figures 12(a)–(e) show the comparison between the TACs of the HDCT image sequence and the reconstructed image sequence via every algorithm. The TACs were extracted for the three 3×3 ROIs located in the region of aorta, HM and IM, respectively, as indicated in figure (f). According to the results, the PWLS-based compared methods tends to overestimate the contrast enhancement of the aorta and the RED-CNN gives the TACs of the myocardia not comparable to the HDCT results. There is no doubt that the TFP-SR method gives the most accurate temporal evolution of the contrast agent referred to the HDCT TACs. Figure 13 shows the MBFs of the porcine heart calculated from the reconstructed DMP-CT images of different algorithms and the reference HDCT images. The advantage of the proposed TFP-SR in low-dose sparse DMP-CT imaging is further revealed. The blood flow values given by TFP-SR in figure 13(e) has visibly higher reference value for clinical diagnosis compared to figures (a)–(d) given by the PWLS-base methods and the RED-CNN, based on the comparison to that of the HDCT images. The MBF seems flattened to some extent, however, the decisive blood flow value was preserved very well.

3.3. Convergence of the TFP-SR algorithm

To state that the proposed TFP-SR method is iteratively convergent, we calculated the relative residual sum (RRS) of the resulting images for every iteration in the digital XCAT phantom experiment. The RRS is defined by



$$\text{RRS}(n) = \frac{\sum_{j=1}^N |\mathbf{x}_j^{\text{true}} - \mathbf{x}_j^n|}{\sum_{j=1}^N |\mathbf{x}_j^{\text{true}} - \mathbf{x}_j^1|}, \quad (30)$$

with n denoting the iteration index. The calculated RRS curves for frames #5, #10, #15 and #24 are shown in figure 14. The convergence of the proposed TFP-SR method can be guaranteed since after enough iterations, the RRS between the reconstructed image and the ground truth images tends to converge to stable values.

4. Conclusion

In this paper, we proposed a TFP-SR method for low-dose DMP-CT image reconstruction. The proposed method involves three key processes:

- Under-sampling and subtraction of the normal-dose baseline sinogram from the low-dose dynamic perfusion sinogram to yield the low-dose enhancement sinogram.

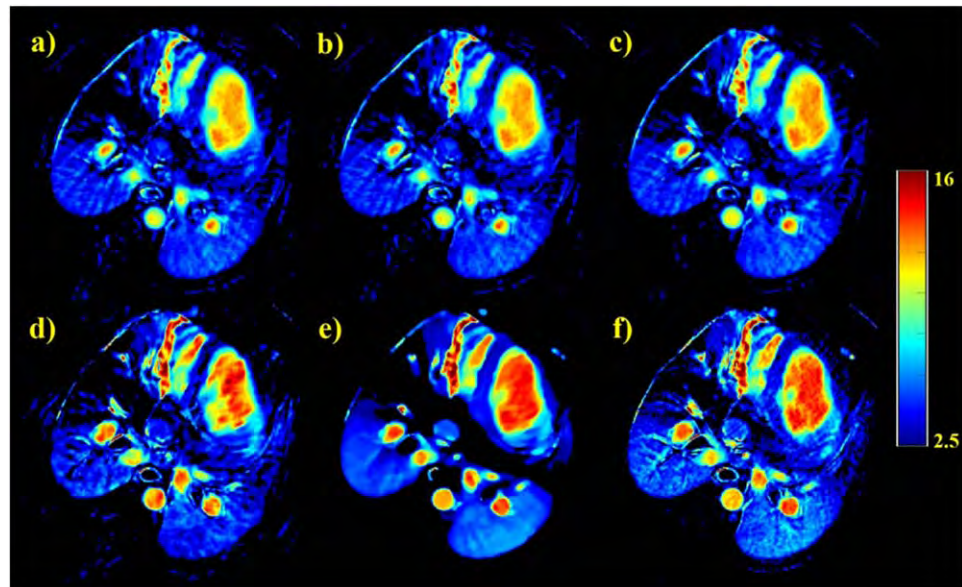


Figure 13. The calculated MBF maps of the porcine DMP-CT image sequences reconstructed by different algorithms. (a) PWLS-TV; (b) PICCS; (c) PWLS-ndiSTV; (d) RED-CNN; (e) TFP-SR; (f) HDCT images.

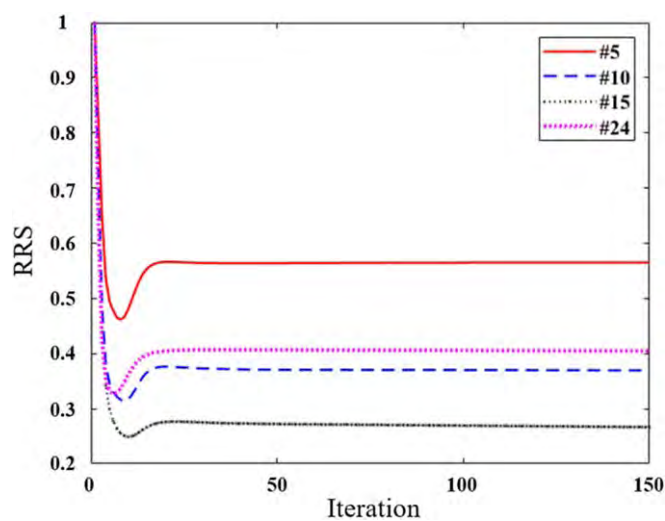


Figure 14. RRS curves calculated using the proposed TFP-SR method.

b. Temporal feature extraction from the pre-reconstructed enhancement image sequence and prior matrix construction.

c. The temporal feature prior-aided reconstruction of the enhancement image sequence.

A dynamic perfusion thoracic XCAT phantom and a sequence of clinical porcine cardiac dynamic perfusion CT images were used to validate the effectiveness and advantages of the proposed TFP-SR method. The experimental results of both the XCAT phantom study and the porcine clinical images study showed that the TFP-SR method has superior performance in terms of imaging visual effect, imaging accuracy and DMP-CT imaging hemodynamics evaluation compared with iterative methods including PWLS-TV, PICCS and PWLS-ndiSTV, and an advanced deep learning network, RED-CNN. With low-dose sparse data acquisition, the advantages of the proposed TFP-SR method are very evident in that the image artifacts that are commonly caused by data incompleteness were successfully eliminated from the reconstructed images, which results in

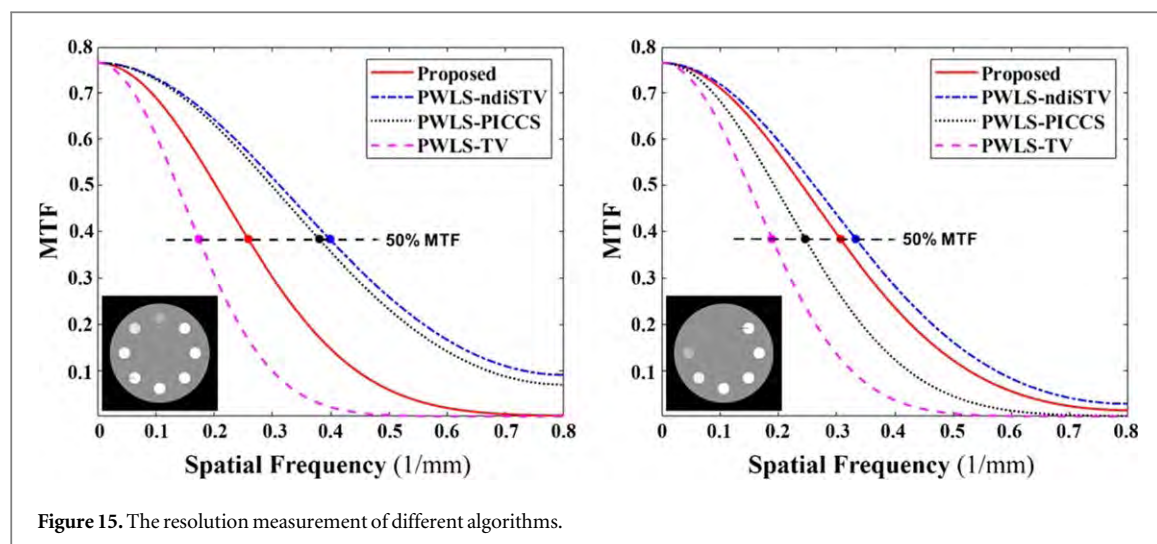


Figure 15. The resolution measurement of different algorithms.

high imaging quality. The concept of separated DMP-CT image reconstruction plays an important role in the success of the proposed method since the structural information from the normal-dose baseline CT image is fully employed, and the advantages of the use of the temporal feature prior are concentrated on the reconstruction of the enhancement images. We conclude that the temporal evolution information naturally possessed by the dynamic image sequence is beneficial for image reconstruction from incomplete projection data, and the methods of temporal information extraction and utilization involved in the proposed TFP-SR algorithm are effective. However, the effectiveness of the proposed TFP-SR method should be further stated based on clinically acquired DMP-CT data before it can be further utilized for clinical DMP-CT imaging.

5. Comments

5.1. Imaging dimension

In clinical medical CT imaging, data acquisition and image reconstruction are based on a cone-beam 3D CT imaging system. However, limited by the lack of effective GPU acceleration computational equipment, in this study, we simulated a monochromatic fan-beam CT imaging system for DMP-CT, for which data acquisition and image reconstruction were conducted based on 2D digital phantoms and clinical images. The proposed TFP-SR method for DMP-CT is validated but must be generalized to obtain an effective 3D imaging system before clinical applications.

5.2. Image misregistration

In our simulation experiments conducted on the digital XCAT thoracic phantom, we assumed strict image registration between the baseline object and the sequential dynamic perfusion objects, which may not be the case in the process of clinical DMP-CT imaging. In the porcine clinical image study, since the simulated HDCT images are acquired from a clinical DMP-CT imaging process, image misregistration, which may result from patient movement, heart motion or z-direction localization error, etc, was considered, from which we may conclude that the proposed TFP-SR method is robust even if image misregistration exists. The ability of the proposed TFP-SR method for dealing with image misregistration comes from the process of sinogram subtraction, by which the difference between the baseline sinogram $^{nd}_b \mathbf{y}$ and the dynamic perfusion sinogram $^{ld}_d \mathbf{y}$ is expressed in the result enhancement sinogram $^{ld}_e \mathbf{y}$, as well as the reconstructed enhancement images $^{ld}_e \mathbf{X}$. By adding the reconstructed enhancement images back to the baseline image, the image misregistration is somewhat compensated.

5.3. Image resolution

The measurement of the resolution of images are necessary for the comprehensive evaluation of an imaging system. However, there are no existed methods for measuring the performance of a CT image reconstruction algorithm in the term of imaging resolution based directly on the reconstructed images. In our study, we used a digital dynamic perfusion resolution phantom to measure the performance of the proposed TFP-SR method. Simulative experiments were conducted using PWLS-TV, PICCS, PWLS-ndiSTV and the proposed TFP-SR methods on the resolution phantom, two image edge profiles (ESF), as marked in the small schematic images in figure 15, were extracted, scaled and fitted to an error function (ERF). The one-dimensional fast Fourier

transformation of the first derivative of the fitted ERF gave the simulative modular transfer function (MTF) with increasing spatial frequency (MTF– f curve). The performance of imaging resolution is indicated by the half-MTF special frequency f_c . According to figure 15, the proposed TFP-SR method does not have the best performance in the term of image resolution. This is because of the natural characteristic of the proposed TFP-SR as it smooths a reconstructed face twice: The effective implementation of the TFP-SR method requires a sequence of noise-controlled pre-reconstructed images, and therefore, the MLEM-TV method was applied. As mentioned in section 2, the involvement of the prior matrix in the iterative reconstruction is closely related to the process of image denoising via a nonlocal method, by which the image is further denoised. However, taking the experimental results in section 3, which tells that the TFP-SR method does not over-smooth the objective image or blur the edges, into consideration, we note that the limitation in term of imaging resolution stated by the resolution measurement experiment is a minor factor for the final imaging quality compared to the image artifact resulted from data incompleteness. The ability of artifact inhibition is more decisive under the condition of ultra-sparse-view DMP-CT imaging.

5.4. Clinical validation

Because clinical raw data are not available in the current study, we validated the proposed TFP-SR method based only on two simulated experiments. The clinical porcine image study is closer to an *in vivo* experiment, from which we may expect that the proposed method will have successful performance in clinical DMP-CT imaging tasks. Further application on clinical DMP-CT reconstruction will be conducted once the desired clinical data are available.

Acknowledgments

The authors would like to thank the editor and anonymous reviewers for their constructive comments and suggestions. This work was supported by the National Natural Science Foundation of China (81871441, 32022042), the Natural Science Foundation of Guangdong Province in China (2020A1515010733), the Guangdong International Science and Technology Cooperation Project of China (2018A050506064), the Shenzhen International Cooperation Research Project of China (GJHZ20180928115824168), the Guangdong Special Support Program of China (2017TQ04R395) and the Chinese Academy of Sciences Key Laboratory of Health Informatics in China (2011DP173015).

Conflict of interest

We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Appendix

A.1. Phantom TAC design

The design of phantom TACs is based on the theories of the human cardiac physiological driving model (Bindschadler *et al* 2014). We assumed a cardiac output of 1.7 l min^{-1} and an MBF level of $3 \text{ ml}/(\text{min} \cdot \text{g tissue})$. The shape of the TAC induced by a single-pass bolus in the cardiac system is often well represented by a simplified gamma variation function (GVF) (Madsen 1992, Mischi *et al* 2008):

$$\gamma(t) = \gamma_{\text{peak}} \left(\frac{t}{t_{\text{peak}}} \right)^{\beta} \exp \left(\beta - \beta \left(\frac{t}{t_{\text{peak}}} \right) \right), \quad (\text{A.1})$$

where γ_{peak} is the maximum value of the function, t_{peak} is the corresponding time point, and β is a parameter controlling the shape of the result GVF. The design of the simulative TACs for the thoracic phantom is conducted by applying different maximum contrast agent concentrations C_{peak} and corresponding max contrast-enhanced time points t_{peak} to different tissues that are investigated. In addition, the arriving time points for different tissues deviate from each other, for which we need to apply a delay time t_{delay} for a particular tissue to express how much the arrival of the contrast agent is delayed compared to the input function, which was defined as the TAC of the aorta. Therefore, the design of the phantom TAC in our work is based on the following expression:

Table A.1. Parameters used for the TAC design.

	C_{peak} (HU)	t_{delay} (s)	t_{peak} (s)	β
AT	90	0.0	5.0	2.0
RV	90	3.0	10.0	2.2
LA	78	6.0	16.0	3.8
LV	75	8.5	17.5	3.2
HM	20	10.5	23.5	1.5
IM	12	10.5	24.5	1.5

$$C(t) = \begin{cases} 0 & \text{if } t < t_{\text{delay}} \\ C_{\text{peak}} \left(\frac{t - t_{\text{delay}}}{t_{\text{peak}}} \right)^{\beta} \exp \left(\beta - \beta \left(\frac{t - t_{\text{delay}}}{t_{\text{peak}}} \right) \right) & \text{if } t > t_{\text{delay}} \end{cases}. \quad (\text{A.2})$$

The cardiac tissues under consideration in our simulation experiments are the aorta (AT), the RV, the left atrium (LA), the LV, the HM and the IM. The parameters for the GVF's for these tissues are given in table 2.

A.2. bSVD method for hemodynamic parameter map calculation

The concentration of contrast agent in a particular cardiac tissue is the response to the biological input function of the contrast agent. Assuming that the TAC of the aorta in this study is the input function of the contrast agent evolution, which is always the case in DMP-CT, the measured concentration of a tissue, for instance, the LV, is a convoluted reaction function to the aorta TAC based on the convolution model of contrast agent transition (Zierler 2000, Lee 2002). The time curve of the measured concentration can be modeled by

$$C(t) = C_{\text{aorta}}(t) \otimes (\rho \cdot F \cdot r(t)), \quad (\text{A.3})$$

where $C_{\text{aorta}}(t)$ is the aorta input function, ρ is the tissue density, F is the blood flow, and $r(t)$ is the residual function. By defining the blood flow scaled residual function

$$k(t) = \rho \cdot F \cdot r(t), \quad (\text{A.4})$$

(A.3) can be replaced by

$$C(t) = C_{\text{aorta}}(t) \otimes k(t), \quad (\text{A.5})$$

Since the TAC measured from the DMP-CT image sequence is time-discrete with a fixed time interval Δt , (A.5) can be discretized as follows:

$$C(t_j) = \int_0^\infty C_{\text{aorta}}(\tau) k(t_j - \tau) d\tau = \sum_{i=1}^T C_{\text{aorta}}(t_i) k(t_{j-i+1}) \Delta t, \quad (\text{A.6})$$

and (A.6) can be written in matrix form as

$$C = Ak. \quad (\text{A.7})$$

With the measured C and the built matrix A , one can solve for the flow scaled residual function k using the singular value decomposition (SVD) (Østergaard *et al* 1996) of matrix A and further estimate the value of the hemodynamic parameters. In practice, the measured aorta input function, namely, $C_{\text{aorta}}'(t)$, may lag the actual aorta input function $C_{\text{aorta}}(t)$ by a delay t_d , which causes the calculated scaled residual function to be not $k(t)$ but $k(t + t_d)$. Wu *et al* proposed the block-circulant SVD method (bSVD) to address this issue in the task of calculating the MBF map from MR perfusion-weighted images (Wu *et al* 2003). This method is applied in this work for the calculation of hemodynamic parameter maps from the reconstructed DMP-CT image sequence. The measured aorta TAC and the tissue TAC, acting as the contrast agent input and output functions, respectively, were zero-padded from length T to length N , where $N = 2T$, to avoid the lag effect. Accordingly, the matrix A in (A.7) is replaced by a block-circulant matrix A^{bc} , of which the element $a_{i,j}^{\text{bc}}$ was calculated by

$$a_{i,j}^{\text{bc}} = \begin{cases} \Delta t C_{\text{aorta}}(t_{i-j+1}) & \text{if } j \leq i \\ \Delta t C_{\text{aorta}}(t_{i-j+1+T}) & \text{if } j > i \end{cases}. \quad (\text{A.8})$$

A^{bc} was then decomposed by $A^{\text{bc}} = U \Sigma V^T$ using the SVD theory, and the estimated flow scaled function k could be restored by

$$k = V \Sigma U^T C. \quad (\text{A.9})$$

Finally, the hemodynamic parameters, including the MBF, MBV and the MTT, could be estimated from $k(t)$ by

$$\text{MBF} = \frac{1}{\rho} \cdot \max(k(t)), \quad (\text{A.10})$$

$$\text{MBV} = \frac{\Delta t}{\rho} \cdot \sum_{j=1}^T k(t_j) \quad (\text{A.11})$$

and

$$\text{MTT} = 60 \times \frac{\text{MBV}}{\text{MBF}}. \quad (\text{A.12})$$

ORCID iDs

Zixiang Chen  <https://orcid.org/0000-0002-9042-4477>

Dong Zeng  <https://orcid.org/0000-0001-6015-5010>

Jianhua Ma  <https://orcid.org/0000-0003-2958-1710>

Dong Liang  <https://orcid.org/0000-0003-1358-9777>

Zhanli Hu  <https://orcid.org/0000-0003-0618-6240>

References

- Bamberg F, Becker A, Schwarz F, Marcus R P, Greif M, von Ziegler F, Blankstein R, Hoffmann U, Sommer W H and Hoffmann V S 2011 Detection of hemodynamically significant coronary artery stenosis: incremental diagnostic value of dynamic CT-based myocardial perfusion imaging *Radiology* **260** 689–98
- Becker A and Becker C 2013 CT imaging of myocardial perfusion: possibilities and perspectives *J. Nucl. Cardiol.* **20** 289–96
- Bian Z, Zeng D, Zhang Z, Gong C, Tian X, Yan G, Huang J, Guo H, Chen B and Zhang J 2017 Low-dose dynamic myocardial perfusion CT imaging using a motion adaptive sparsity prior *Med. Phys.* **44** e188–201
- Bindschadler M, Modgil D, Branch K R, La Riviere P J and Alessio A M 2014 Comparison of blood flow models and acquisitions for quantitative myocardial perfusion estimation from dynamic CT *Phys. Med. Biol.* **59** 1533–56
- Chávez-Rivera L B, Ortega-Máynez L, Mejía J and Mederos B 2015 *IEEE Nuclear Science Symp. and Medical Imaging Conf. (NSS/MIC)*, 2015 (Piscataway, NJ: IEEE) pp 1–5
- 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
- Chen H, Zhang Y, Kalra M K, Lin F, Chen Y, Liao P, Zhou J and Wang G 2017 Low-dose CT with a residual encoder-decoder convolutional neural network *IEEE Trans. Med. Imaging* **36** 2524–35
- Donoho D L 2006 Compressed sensing *Trans. Inf. Theory* **52** 1289–306
- Enjilela E, Lee T-Y, Hsieh J, Wisenberg G, Teefy P, Yadegari A, Bagur R, Islam A, Branch K and So A 2018 Ultra-low dose quantitative CT myocardial perfusion imaging with sparse-view dynamic acquisition and image reconstruction: a feasibility study *Int. J. Cardiol.* **254** 272–81
- Friedman J H, Bentley J L and Finkel R A 1977 An algorithm for finding best matches in logarithmic expected time *ACM Trans. Math. Softw.* **3** 209–26
- Geyer L L *et al* 2015 State of the art: iterative CT reconstruction techniques *Radiology* **276** 339–57
- Gong C, Han C, Gan G, Deng Z, Zhou Y, Yi J, Zheng X, Xie C and Jin X 2017 Low-dose dynamic myocardial perfusion CT image reconstruction using pre-contrast normal-dose CT scan induced structure tensor total variation regularization *Phys. Med. Biol.* **62** 2612–35
- Ho K-T, Chua K-C, Klotz E and Panknin C 2010 Stress and rest dynamic myocardial perfusion imaging by evaluation of complete time-attenuation curves with dual-source CT *JACC Cardiovascular Imaging* **3** 811–20
- Hofmann T, Schölkopf B and Smola A J 2008 Kernel methods in machine learning *Ann. Stat.* **1171**–220
- Hore A and Ziou D 2010 *20th Int. Conf. on Pattern Recognition, 2010* (Piscataway, NJ: IEEE) pp 2366–9
- Hu Z, Jiang C, Sun F, Zhang Q, Ge Y, Yang Y, Liu X, Zheng H and Liang D 2019 Artifact correction in low-dose dental CT imaging using Wasserstein generative adversarial networks *Med. Phys.* **46** 1686–96
- Hu Z, Zhang Y, Liu J, Ma J, Zheng H and Liang D 2016 A feature refinement approach for statistical interior CT reconstruction *Phys. Med. Biol.* **61** 5311–34
- Huang J, Ma J, Liu N, Zhang H, Bian Z, Feng Y, Feng Q and Chen W 2011 Sparse angular CT reconstruction using non-local means based iterative-correction POCS *Comput. Biol. Med.* **41** 195–205
- Kang E, Min J and Ye J C 2017 A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction *Med. Phys.* **44** e360–75
- Kim S M, Kim Y N and Choe Y H 2013 Adenosine-stress dynamic myocardial perfusion imaging using 128-slice dual-source CT: optimization of the CT protocol to reduce the radiation dose *Int. J. Cardiovascular Imaging* **29** 875–84
- Lange K and Carson R 1984 EM reconstruction algorithms for emission and transmission tomography *J. Comput. Assist. Tomogr.* **8** 306–16
- Lee T-Y 2002 Functional CT: physiological models *Trends Biotechnol.* **20** S3–10
- Li Y, Speidel M A, François C J and Chen G-H 2017 Radiation dose reduction in CT myocardial perfusion imaging using SMART-RECON *IEEE Trans. Med. Imaging* **36** 2557–68
- Ma J, Zhang H, Gao Y, Huang J, Liang Z, Feng Q and Chen W 2012 Iterative image reconstruction for cerebral perfusion CT using a pre-contrast scan induced edge-preserving prior *Phys. Med. Biol.* **57** 7519–42
- Madsen M T 1992 A simplified formulation of the gamma variate function *Phys. Med. Biol.* **37** 1597–600
- Mischi M, Den Boer J and Korsten H 2008 On the physical and stochastic representation of an indicator dilution curve as a gamma variate *Physiol. Meas.* **29** 281–94
- Niu S, Gao Y, Bian Z, Huang J, Chen W, Yu G, Liang Z and Ma J 2014 Sparse-view x-ray CT reconstruction via total generalized variation regularization *Phys. Med. Biol.* **59** 2997–3018

- Østergaard L, Weisskoff R M, Chesler D A, Gyldensted C and Rosen B R 1996 High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis *Magn. Reson. Med.* **36** 715–25
- Panin V, Zeng G and Gullberg G 1999 Total variation regulated EM algorithm [SPECT reconstruction] *IEEE Trans. Nucl. Sci.* **46** 2202–10
- Rockmore A J and Macovski A 1976 A maximum likelihood approach to emission image reconstruction from projections *IEEE Trans. Nucl. Sci.* **23** 1428–32
- Shan H, Padole A, Homayounieh F, Kruger U, Khera R D, Nitiwarangkul C, Kalra M K and Wang G 2019 Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose CT image reconstruction *Nat. Mach. Intell.* **1** 269–76
- Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113–22
- Speidel M A, Van Lysel M S, Reeder S B, Supanich M, Nett B E, Zambelli J, Chang S M, Hsieh J, Chen G-H and Mistretta C A 2007 *Medical Imaging 2007: Physics of Medical Imaging* (vol 6510) (Int. Society for Optics and Photonics) p 651014
- Tao Y, Chen G H, Hacker T A, Raval A N, Van Lysel M S and Speidel M A 2014 Low dose dynamic CT myocardial perfusion imaging using a statistical iterative reconstruction method *Med. Phys.* **41** 071914
- Uren N G, Melin J A, De Bruyne B, Wijns W, Baudhuin T and Camici P G 1994 Relation between myocardial blood flow and the severity of coronary-artery stenosis *New Engl. J. Med.* **330** 1782–8
- Wang G, Schweiger G and Vannier M W 1998 An iterative algorithm for x-ray CT fluoroscopy *IEEE Trans. Med. Imaging* **17** 853–6
- Williams M and Newby D 2016 CT myocardial perfusion imaging: current status and future directions *Clin. Radiol.* **71** 739–49
- Wu O, Østergaard L, Weisskoff R M, Benner T, Rosen B R and Sorensen A G 2003 Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix *Magn. Reson. Med.* **50** 164–74
- Xie S, Zheng X, Chen Y, Xie L, Liu J, Zhang Y, Yan J, Zhu H and Hu Y 2018 Artifact removal using improved GoogLeNet for Sparse-view CT reconstruction *Sci. Rep.* **8** 6700
- Zierler K 2000 Indicator dilution methods for measuring blood flow, volume, and other properties of biological systems: a brief history and memoir *Ann. Biomed. Eng.* **28** 836–48