

Regular Article

Low-Dose CT Image Denoising using Image Decomposition and Sparse Representation

Nguyen Thanh Trung^{1,2}, Trinh Dinh Hoan³, Nguyen Linh Trung¹, Marie Luong⁴

¹ Advanced Institute of Engineering and Technology (AVITECH), University of Engineering and Technology, Vietnam National University, Hanoi, Vietnam

² University of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Vietnam

³ Université de Lorraine, CNRS, CRAN, 54000 Nancy, France

⁴ L2TI Laboratory, Galillee Institute, University Paris 13, 93430 Villetaneuse, France

Correspondence: Nguyen Linh Trung, linhtrung@vnu.edu.vn

Communication: received 7 June 2019, revised 20 September 2019, accepted 30 September 2019

Online publication: 23 November 2019, Digital Object Identifier: 10.21553/rev-jec.238

The associate editor coordinating the review of this article and recommending it for publication was Prof. Vo Nguyen Quoc Bao.

Abstract– X-ray computed tomography (CT) is now a widely used imaging modality for numerous medical purposes. The risk of high X-ray radiation may induce genetic, cancerous and other diseases, demanding the development of new image processing methods that are able to enhance the quality of low-dose CT images. However, lowering the radiation dose increases the noise in acquired images and hence affects important diagnostic information. This paper contributes an efficient denoising method for low-dose CT images. A noisy image is decomposed into three component images of low, medium and high frequency bands; noise is mainly presented in the medium and high component images. Then, by exploiting the fact that a small image patch of the noisy image can be approximated by a linear combination of several elements in a given dictionary of noise-free image patches generated from noise-free images taken at nearly the same position with the noisy image, noise in these medium and high component images are effectively eliminated. Specifically, we give new solutions for image decomposition to easily control the filter parameters, for dictionary construction to improve the effectiveness and reduce the running-time. Instead of using a large dataset of patches, only a structured small part of patches extracted from the raw data is used to form a dictionary, to be used in sparse coding. In addition, we illustrate the effectiveness of the proposed method in preserving image details which are subtle but clinically important. Experimental results conducted on both synthetic and real noise data demonstrate that the proposed method is competitive with the state-of-the-art methods.

Keywords– Computed Tomography (CT), medical image, low-dose radiation, patch-based image denoising, image decomposition, sparse representation.

1 INTRODUCTION

Computed Tomography (CT), also called computerized axial tomography, is one of the most important medical imaging techniques, and uses X-rays to create cross-sectional images of the body. CT images are used for diagnostic and therapeutic purposes. However, a great concern to patients and operators is the risk of high X-ray radiation which may induce genetic, cancerous and other diseases. Therefore, it is important to reduce the radiation dose as much as possible while preserving the image quality for clinical purposes. This calls for major effort in the CT research community for developing new image processing methods that are able to enhance the quality of low-dose CT images. However, lowering the radiation dose increases the noise in acquired images, as illustrated in Figure 1, and hence affects important diagnostic information. How to denoise low-dose CT images such that the quality of denoised images is as close as possible to that of normal-dose CT images is the concern of this paper.

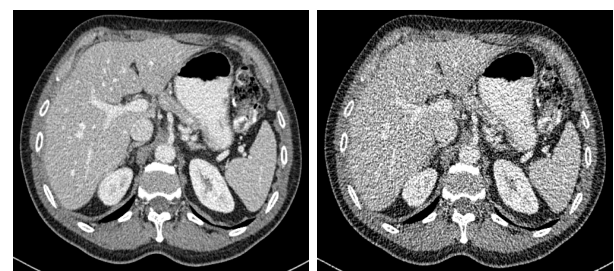


Figure 1. CT images of the liver at the same position with normal and low radiation doses.

Numerous denoising methods have been proposed in the literature. The classical noise filters such as the Gaussian filter [1], the Wiener filter [2] and the bilateral filter [3] could effectively reduce noise in homogeneous regions but often suppress high frequency structures such as edges or subtle details.

To overcome this drawback of the classical noise filters, numerous other denoising approaches were pro-

posed, such as the total variation (TV/TGV) based methods [4–7], the non-local means (NLM) based methods [8–10], the sparse representation based methods [11–14], and the 3D block-matching (BM3D) based filters [15–17]. However, subtle details are still over-smoothed. Thus, it is difficult to directly apply these methods to medical images.

Example learning based image processing, as earlier seen in for example [18], and has recently become an attractive approach for CT image denoising as seen in [19–21] using convolutional neural networks (CNN). Given a training set of image pairs $\{(x_i, y_i)\}$ where x_i is a noisy image of size $m \times m$ and y_i is the corresponding noise-free (clean) image of size $n \times n$ ($1 \leq n \leq m$), the main idea of learning methods is to estimate the map from the space of the noisy image to the space of the noise-free image. In [19–21], the training dataset is established from normal-dose and low-dose CT image pairs. It is shown that CNN-based methods can effectively reduce noise in low-dose CT images. Generally, the performance of a learning-based denoising method highly depends on the quality of the training dataset and the method to establish the training data.

Another learning-based approach for CT image denoising is based on sparse representation [22–25]. In this approach, a dictionary of noise-free image patches is first created from a given set of standard example CT images; by “standard”, we mean that these example images are either noiseless or of high quality. Denoising a noisy image is then performed patch-wise by estimating the sparse representation of each patch of this noisy image using the dictionary. Although denoising of low-dose CT images by learning-based methods has obtained significant achievements, preserving subtle details which contain important pathological information remains a challenge.

We mention here two other works in [26, 27], based on which we develop further in this paper. In order to preserve subtle details in the denoised image, the main idea in [26] was to decompose an image into component images corresponding to three frequency bands: low, medium and high. The low and high frequency images can be easily obtained by any traditional denoising method. Denoising in [26] focuses on estimating the high frequency image. To this end, an example dataset of noise-free medium-high patch pairs is built from a given set of standard images. Markov random field (MRF) is then used to find in this example dataset the best candidate of the high frequency component patch in a noisy patch. By that way, it was demonstrated that noise can be effectively removed while small details in the image were well preserved. However, since the estimated high frequency component image was directly synthesized from the dataset of patch-pairs, the performance of the method highly depends on the example images. Moreover, this method is time-consuming.

A more effective denoising strategy was proposed in [27] where sparse representation was used, instead of the MRF, to estimate both the high and medium frequency component patches of a noisy patch. To

denoise a patch x , a sub-dictionary including medium-high frequency patch pairs, in which the medium frequency patches are neighbors of the medium frequency component of x in the example dataset. The sparse representation of the medium frequency component of x over the example medium frequency patches in the sub-dictionary helps effectively estimate the high frequency component of x . The efficiency of this method comes from the fact that, for each patch, denoising is based on an adaptive example sub-dictionary. However, similar to the method in [26], a drawback of this method is also time-consuming.

In both [26, 27], the decomposition of an image (noisy or example) into component images of different frequency bands is performed by two Gaussian low-pass filters. The parameters of one filter, in terms of size and variance, depend on the other and thus make it difficult to set the filter parameters.

To improve on [26, 27], this paper proposes a new patch-based denoising method with three contributions. First, we propose a new solution for image decomposition that is easier to control the filter parameters. Second, we propose a new solution for dictionary construction to improve the effectiveness and reduce the computational complexity of the method in terms of running-time. Instead of using a large dataset of patches, only a structured small part of patches extracted from the raw data is used to form a dictionary, to be used in sparse coding. Third, we illustrate the effectiveness of the proposed method in preserving image details which are subtle but clinically important, in a particular case of a cancerous nodule in a lung image of a cancer patient. The proposed method is competitive with the state-of-the-art methods.

The paper is organized as follows. After presenting some background information in Section 2, we present the proposed method in Section 3. The performance evaluation of the proposed method is shown in Section 4. Finally, conclusions are given in Section 5.

2 BACKGROUND

Usually, in order to denoise effectively, denoising methods have to be designed based on the distribution of noise. As shown in [21, 22], the noise distribution in a CT image is not only non-uniform but also complex over the whole image. For low-dose CT images, this complexity of the noise model leads to poor performance when using the traditional Gaussian and/or Poisson prior-based methods. Trinh *et al.* in [22, 23] proposed a useful assumption that the noise distribution in CT images can be locally approximated by a zero-mean Gaussian distribution. It is noteworthy that, on the same image, noise levels may be different for different positions. This assumption enables patch-based denoising methods to use the Gaussian noise assumption on small image patches.

Let Y be a noisy low-dose CT image. The objective of this study is to estimate its ideal noise-free version X . We use the assumption about the noise as that given

in [22], that is, if \mathbf{y}_i centered at pixel i is a small patch of \mathbf{Y} , then

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{n}_i, \quad (1)$$

where \mathbf{x}_i is the corresponding noise-free patch in \mathbf{X} , \mathbf{n}_i is additive white zero-mean Gaussian noise with variance of σ_i^2 . Given \mathbf{y}_i , we want to estimate \mathbf{x}_i . In this paper, we are interested in using sparse coding for patch-based denoising.

The goal of sparse coding is to find a sparse representation of a vector $\mathbf{x} \in \mathbb{R}^d$ over a given set of sample vectors $\{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N\} \subset \mathbb{R}^d$. This set can be rewritten in matrix form as $\mathbf{D} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N]$. Normally, it is assumed that $d < N$, \mathbf{c}_i ($i = 1, 2, \dots, N$) are ℓ_2 -normalized vectors, and \mathbf{D} is considered as a dictionary. Then, if there exists a vector $\boldsymbol{\alpha} \in \mathbb{R}^N$ containing a very few non-zero entries such that $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$, then the sparse vector $\boldsymbol{\alpha}$ is often estimated by solving the following optimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \quad \text{s.t.} \quad \|\boldsymbol{\alpha}\|_0 \leq L, \quad (2)$$

or

$$\hat{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \quad \text{s.t.} \quad \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 \leq \varepsilon, \quad (3)$$

where $\|\boldsymbol{\alpha}\|_0$ is the ℓ_0 -pseudo-norm, which counts the non-zero entries in $\boldsymbol{\alpha}$, L is the maximum number of non-zero elements in $\boldsymbol{\alpha}$ ($L \ll N$), and ε is a tolerance parameter.

Numerous sparse-coding-based denoising methods have been proposed; the idea comes from [12]. Elad *et al.* in [12] introduced an effective denoising method, referred to as KSVD. The authors used noisy patches \mathbf{y}_i extracted from image \mathbf{Y} as data for training a dictionary \mathbf{D} , to be then used in sparse-coding. \mathbf{D} is determined by solving the following optimization problem:

$$\min_{\mathbf{D}, \boldsymbol{\alpha}_i} \sum_i \|\mathbf{y}_i - \mathbf{D}\boldsymbol{\alpha}_i\|_2^2 + \lambda \|\boldsymbol{\alpha}_i\|_p \quad \text{s.t.} \quad \|\mathbf{D}(:, k)\|_2 = 1,$$

where λ is a trade-off parameter controlling the sparsity penalty and the representation fidelity, $\|\mathbf{D}(:, k)\|_2$ denotes the k -th column of \mathbf{D} , and $\|\bullet\|_p$ is ℓ_p -norm ($0 \leq p \leq 1$). After training the dictionary \mathbf{D} , $\mathbf{D}\boldsymbol{\alpha}_i$ is considered as the denoised version of \mathbf{y}_i . Although KSVD can effectively remove noise, subtle details in the image are often over-smoothed.

3 PROPOSED METHOD

Suppose that we need to restore an ideal CT image \mathbf{X} from its noisy low-dose image \mathbf{Y} satisfying (1) with the help of a given set $\Omega = \{\mathbf{I}_t\}_{t=1}^T$ of T normal-dose images, which are considered as noise-free images and taken at nearly the same position as \mathbf{Y} .

The method in [26] proposed a reasonable assumption that if \mathbf{Y} is decomposed into three component images corresponding to three frequency bands (low, medium and high), as

$$\mathbf{Y} = \mathbf{Y}^{\text{low}} + \mathbf{Y}^{\text{mid}} + \mathbf{Y}^{\text{high}}, \quad (4)$$

then the majority of the noise is included in \mathbf{Y}^{high} and the rest in \mathbf{Y}^{mid} . Hence, to denoise \mathbf{Y} , we need

to estimate the noise-free component images \mathbf{X}^{mid} and \mathbf{X}^{high} of \mathbf{Y}^{mid} and \mathbf{Y}^{high} , respectively. Finally,

$$\hat{\mathbf{X}} = \mathbf{Y}^{\text{low}} + \mathbf{X}^{\text{mid}} + \mathbf{X}^{\text{high}} \quad (5)$$

is the denoising result (an estimate of \mathbf{X}).

The method proposed in this paper follows the above idea of image decomposition. The main blocks of the proposed method are image decomposition, database construction, sparse-coding-based denoising, and image composition, as shown in Figure 2 and presented next.

3.1 Image Decomposition

According to [26, 27], an image \mathbf{I} , whether being noise-free or noisy, is decomposed into three component images \mathbf{I}^{low} , \mathbf{I}^{mid} and \mathbf{I}^{high} , using two Gaussian filters $\mathcal{F}_1^{\text{low}}$ and $\mathcal{F}_2^{\text{low}}$. Specifically,

$$\begin{aligned} \mathbf{I}^{\text{low}} &= \mathcal{F}_1^{\text{low}} \mathbf{I}, \\ \mathbf{I}^{\text{mid}} &= \mathcal{F}_2^{\text{low}} \mathbf{I} = \mathcal{F}_2^{\text{low}} \mathbf{I} - \mathcal{F}_1^{\text{low}} \mathbf{I}, \\ \mathbf{I}^{\text{high}} &= \mathbf{I} - \mathbf{I}^{\text{mid}} - \mathbf{I}^{\text{low}}. \end{aligned}$$

As mentioned in Section 1, to properly obtain \mathbf{I}^{mid} of \mathbf{I} , the parameters of $\mathcal{F}_2^{\text{low}}$, in terms of size and variance, depend on those of $\mathcal{F}_1^{\text{low}}$, and thus make it difficult to set the filter parameters. For example, if we mistakenly set the size and the variance of $\mathcal{F}_2^{\text{low}}$ to be the same as $\mathcal{F}_1^{\text{low}}$, then \mathbf{I}^{mid} is null.

To overcome this difficulty, we propose in this paper a new solution of decomposition, as follows:

$$\mathbf{I}^{\text{high}} = \mathbf{I} - \mathcal{F}_1^{\text{low}} \mathbf{I}, \quad (6)$$

$$\mathbf{I}^{\text{low}} = \mathcal{F}_2^{\text{low}} \mathcal{F}_1^{\text{low}} \mathbf{I}, \quad (7)$$

$$\mathbf{I}^{\text{mid}} = \mathcal{F}_1^{\text{low}} \mathbf{I} - \mathbf{I}^{\text{low}} = \mathcal{F}_1^{\text{low}} \mathbf{I} - \mathcal{F}_2^{\text{low}} \mathcal{F}_1^{\text{low}} \mathbf{I}. \quad (8)$$

We can see that this decomposition method also satisfies condition (4), $\mathbf{I} = \mathbf{I}^{\text{high}} + \mathbf{I}^{\text{mid}} + \mathbf{I}^{\text{low}}$. Equation (8) shows that the frequency band of \mathbf{I}^{mid} is lower than that of \mathbf{I}^{high} and higher than that of \mathbf{I}^{low} . In addition, the dependency of parameter setting of $\mathcal{F}_2^{\text{low}}$ on $\mathcal{F}_1^{\text{low}}$ is reduced; for example, we can set them to have the same size and variance.

3.2 Dictionary Construction

Similar to [27], in this paper denoising is performed patch-wise on \mathbf{Y}^{mid} and \mathbf{Y}^{high} . A database of medium-high frequency image patch pairs is constructed from a set Ω of standard images \mathbf{I}_t , $t = 1, \dots, T$.

Each example image \mathbf{I}_t is decomposed into three component images $\mathbf{I}_t^{\text{low}}$, $\mathbf{I}_t^{\text{mid}}$ and $\mathbf{I}_t^{\text{high}}$, according to (6), (7) and (8). For each image pair $(\mathbf{I}_t^{\text{mid}}, \mathbf{I}_t^{\text{high}})$, a set of patch pairs is generated by randomly extracting $(\sqrt{n} \times \sqrt{n})$ -patches from $\mathbf{I}_t^{\text{mid}}$ and $\mathbf{I}_t^{\text{high}}$. These selected patches are vectorized and scaled to obtain a sample dataset \mathbf{D}_t , as given by

$$\mathbf{D}_t = \left\{ (\mathbf{c}_i^{\text{mid},t}, \mathbf{c}_i^{\text{high},t}) = \left(\frac{\mathbf{p}_i^{\text{mid},t}}{\|\mathbf{p}_i^{\text{mid},t}\|}, \frac{\mathbf{p}_i^{\text{high},t}}{\|\mathbf{p}_i^{\text{high},t}\|} \right) \right\} \subset \mathbb{R}^n \times \mathbb{R}^n, \quad (9)$$

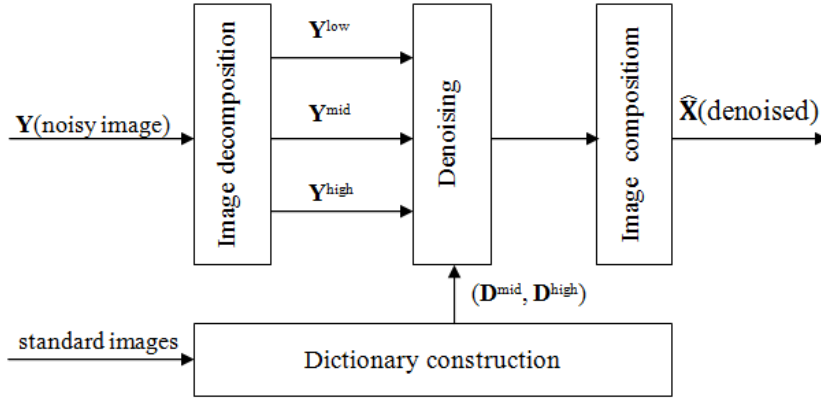


Figure 2. Diagram of the proposed method.

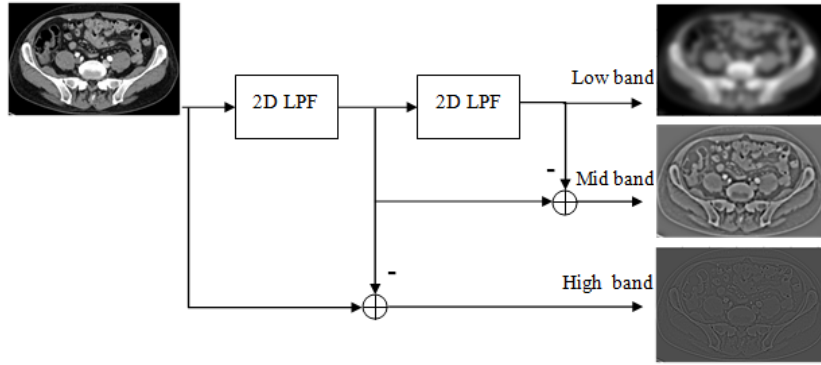


Figure 3. Image decomposition using 2D low-pass filters.

where $\mathbf{p}_i^{\text{mid},t}$ and $\mathbf{p}_i^{\text{high},t}$ correspond to the $(\sqrt{n} \times \sqrt{n})$ -patches at pixel i in images $\mathbf{I}_t^{\text{mid}}$ and $\mathbf{I}_t^{\text{high}}$, respectively.

Then, an overall database of normalized vector pairs is synthesized from T datasets \mathbf{D}_t , as given by

$$\Sigma = \bigcup_{t=1}^T \mathbf{D}_t = \{(\mathbf{c}_i^{\text{mid}}, \mathbf{c}_i^{\text{high}})\}_{i=1}^{N_\Omega}.$$

Due to the spatial redundancy on image patches in the standard images $\mathbf{I}_t \in \Omega$, there exist a large number of similar elements in the overall database Σ . It would therefore be time-consuming if sparse coding is performed on Σ . To deal with this issue, we propose in this paper a solution for data reduction, to create a dictionary of size much smaller than that of the database.

Consider two normalized vectors $\mathbf{c}_i^{\text{mid}}$ and $\mathbf{c}_j^{\text{mid}}$ in Σ , we have

$$\|\mathbf{c}_i^{\text{mid}} - \mathbf{c}_j^{\text{mid}}\|^2 = 2(1 - \mathbf{c}_i^{\text{mid}} \cdot \mathbf{c}_j^{\text{mid}}).$$

Thus, $\mathbf{c}_i^{\text{mid}}$ and $\mathbf{c}_j^{\text{mid}}$ are called ϵ -similar if their scalar product $\mathbf{c}_i^{\text{mid}} \cdot \mathbf{c}_j^{\text{mid}} \geq \epsilon$ with ϵ being close to 1. Therefore, to reduce Σ , we only keep one vector among ϵ -similar vectors. Finally, a dictionary \mathbf{D} for sparse coding is obtained

$$\mathbf{D} = (\mathbf{D}^{\text{mid}}, \mathbf{D}^{\text{high}}) = \{(\mathbf{c}_i^{\text{mid}}, \mathbf{c}_i^{\text{high}})\}_{i=1}^N, \quad (10)$$

where $\mathbf{D}^{\text{mid}} = \{\mathbf{c}_i^{\text{mid}}\}_{i=1}^N$ and $\mathbf{D}^{\text{high}} = \{\mathbf{c}_i^{\text{high}}\}_{i=1}^N$ such that

$$\mathbf{c}_i^{\text{mid}} \cdot \mathbf{c}_j^{\text{mid}} < \epsilon, \forall i \neq j.$$

In our method, ϵ is empirically set to 0.99. The computation time to obtain the dictionary \mathbf{D} of 7092 atoms from set Σ with $N_\Omega = 1522804$ and $n = 25$, as an example, is 40.485 seconds.

3.3 Sparse-coding-based Denoising

Suppose we need to denoise an image \mathbf{Y} satisfying assumption (1). We first decompose it into three component images \mathbf{Y}^{low} , \mathbf{Y}^{mid} , and \mathbf{Y}^{high} , according to (6), (7) and (8). Then, for every patch \mathbf{y}_i in \mathbf{Y} , we have

$$\mathbf{y}_i = \mathbf{y}_i^{\text{low}} + \mathbf{y}_i^{\text{mid}} + \mathbf{y}_i^{\text{high}},$$

where $\mathbf{y}_i^{\text{low}} \in \mathbf{Y}^{\text{low}}$, $\mathbf{y}_i^{\text{mid}} \in \mathbf{Y}^{\text{mid}}$ and $\mathbf{y}_i^{\text{high}} \in \mathbf{Y}^{\text{high}}$.

Following (5), the desired patch \mathbf{x}_i is estimated by

$$\mathbf{x}_i = \mathbf{y}_i^{\text{low}} + \mathbf{x}_i^{\text{mid}} + \mathbf{x}_i^{\text{high}},$$

where $\mathbf{x}_i^{\text{mid}}$ and $\mathbf{x}_i^{\text{high}}$ are medium and high frequency component patches of \mathbf{x}_i , which will be estimated from $\mathbf{y}_i^{\text{mid}}$ and $\mathbf{y}_i^{\text{high}}$, respectively. Following (9) and (10), we have

$$\mathbf{c}_i^{\text{mid}} + \mathbf{c}_i^{\text{high}} = \frac{(\mathbf{p}_i - \mathbf{p}_i^{\text{low}})}{\|\mathbf{p}_i^{\text{mid}}\|},$$

for all $(\mathbf{c}_i^{\text{mid}}, \mathbf{c}_i^{\text{high}}) \in \mathbf{D}$, where \mathbf{p}_i is a vectorized noise-free patch. Thus, corresponding to \mathbf{D} , $\mathbf{G} = \{\mathbf{g}_i = (\mathbf{p}_i - \mathbf{p}_i^{\text{low}})/\|\mathbf{p}_i^{\text{mid}}\|\}_{i=1}^N$ can be considered as a dictionary containing the middle and high frequency

components of \mathbf{p}_i . Consequently, if

$$(\mathbf{x}_i - \mathbf{x}_i^{\text{low}}) = \sum_{i=1}^N \beta_i \mathbf{g}_i$$

is a sparse representation of $(\mathbf{x}_i - \mathbf{x}_i^{\text{low}})$ on \mathbf{G} , then

$$\mathbf{x}_i^{\text{mid}} + \mathbf{x}_i^{\text{high}} = \sum_{i=1}^N \beta_i \mathbf{g}_i = \sum_{i=1}^N (\beta_i \mathbf{c}_i^{\text{mid}}) + \sum_{i=1}^N (\beta_i \mathbf{c}_i^{\text{high}}).$$

Hence, we can consider

$$\begin{aligned} \mathbf{x}_i^{\text{mid}} &= \sum_{i=1}^N (\beta_i \mathbf{c}_i^{\text{mid}}), \\ \mathbf{x}_i^{\text{high}} &= \sum_{i=1}^N (\beta_i \mathbf{c}_i^{\text{high}}). \end{aligned}$$

It means that $\mathbf{x}_i^{\text{mid}}$ and $\mathbf{x}_i^{\text{high}}$ have the same sparse representation on \mathbf{D}^{mid} and \mathbf{D}^{high} . This leads to a reasonable assumption that with \mathbf{x}_i is a patch of size $(\sqrt{n} \times \sqrt{n})$ as the standard patches in database \mathbf{D} vectorized patches $\mathbf{x}_i^{\text{mid}}$ and $\mathbf{x}_i^{\text{high}}$ have the same sparse representation on \mathbf{D}^{mid} and \mathbf{D}^{high} , respectively. Therefore, the corresponding patches $\mathbf{y}_i^{\text{mid}}$ of $\mathbf{x}_i^{\text{mid}}$ and $\mathbf{y}_i^{\text{high}}$ of $\mathbf{x}_i^{\text{high}}$ also have the same sparse representation on \mathbf{D}^{mid} and \mathbf{D}^{high} .

Since $\mathbf{y}_i^{\text{mid}}$ is less noisy than $\mathbf{y}_i^{\text{high}}$, the sparse representation will be determined by sparse-coding of $\mathbf{y}_i^{\text{mid}}$ over \mathbf{D}^{mid} . In this paper, we use the sparse-coding model given in (3), as given by

$$\begin{aligned} \hat{\alpha}_i &= \arg \min_{\alpha_i} \|\alpha_i\|_0, \\ \text{s.t. } &\|\mathbf{D}^{\text{mid}} \alpha_i - \mathbf{y}_i^{\text{mid}}\|_2^2 \leq \gamma(n\sigma_i^2), \end{aligned} \quad (11)$$

where γ is a threshold parameter. Here, γ is not too sensitive with the noise level σ_i (standard deviation) in \mathbf{y}_i because $\mathbf{y}_i^{\text{mid}}$ is a noiseless patch.

Then, $\mathbf{x}_i^{\text{mid}}$ and $\mathbf{x}_i^{\text{high}}$ are estimated as

$$\begin{aligned} \hat{\mathbf{x}}_i^{\text{mid}} &= \mathbf{D}^{\text{mid}} \hat{\alpha}_i, \\ \hat{\mathbf{x}}_i^{\text{high}} &= \mathbf{D}^{\text{high}} \hat{\alpha}_i. \end{aligned}$$

Finally, we obtain the following estimate of the denoised patch:

$$\hat{\mathbf{x}}_i = \mathbf{y}_i^{\text{low}} + \hat{\mathbf{x}}_i^{\text{mid}} + \hat{\mathbf{x}}_i^{\text{high}}.$$

Comparing the sparse-coding model (2) used in [27] and recalled below

$$\begin{aligned} \hat{\alpha} &= \arg \min_{\alpha} \frac{1}{2} \|\mathbf{D}_i^{\text{mid}} \mathbf{f} \mathbf{f} - \mathbf{y}_i^{\text{mid}}\|_2^2, \\ \text{s.t. } &\|\mathbf{f} \mathbf{f}\|_0 \leq L, \|\mathbf{D}_i^{\text{high}} \alpha - \mathbf{y}_i^{\text{high}}\|_2^2 \leq \lambda \sigma_i^2, \end{aligned} \quad (12)$$

where $\mathbf{D}_i^{\text{high}}$ and $\mathbf{D}_i^{\text{mid}}$ are K -atom sub-dictionaries extracted from \mathbf{D}^{mid} and \mathbf{D}^{high} , the sparse-coding model used in this paper is more effective because it is difficult to set an optimal value for the threshold parameter L in (12) while setting values for parameter γ in (11) is adaptive to the noise level of patches. Moreover, the model in (11) uses only one dictionary pair $(\mathbf{D}^{\text{mid}}, \mathbf{D}^{\text{high}})$ for all patches and, therefore, reduces the computational time and does not need the parameter K .

3.4 Image Composition

Having obtained all the estimates $\hat{\mathbf{x}}_i$ of all pixels i in \mathbf{Y} , we then combine them to obtain the final denoised overall image, as shown below, by using the method proposed in [12]:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X}} \eta \|\mathbf{X} - \mathbf{Y}\|_2^2 + \sum_i \|\hat{\mathbf{x}}_i - \mathbf{R}_i \mathbf{X}\|_2^2, \quad (13)$$

where \mathbf{R}_i is a matrix of size $n \times M$ that extracts and then vectorizes a patch of size $\sqrt{n} \times \sqrt{n}$ in an image \mathbf{X} of size $W \times H$; here $M = W \cdot H$.

4 PERFORMANCE EVALUATION

To evaluate the performance of proposed method, we perform experiments on both synthetic and real low-dose CT images. All used images are 8-bit grayscale ones. The proposed method is compared to the state-of-the-art denoising methods, namely non-local means (NLM) [8], total generalized variation (TGV) [28], and KSVD [12]. Moreover, it is also compared to several previous learning-based methods proposed in [26] (referred to as MRFD) and in [27] (referred to as FD-SC1) to clearly see the improvements. The proposed method is referred to as FD-SC2 (this name stands for Frequency Decomposition and Sparse Coding).

In FD-SC2, the Gaussian low-pass filters $\mathcal{F}_1^{\text{low}}$ and $\mathcal{F}_2^{\text{low}}$ has the same size of 7×7 , and the standard deviation of 1 and 3, respectively. The overlap size between two neighborhood $\sqrt{n} \times \sqrt{n}$ -patches is set to $(\sqrt{n} - 1)$. The threshold parameter γ in (11) and the patch-size parameter will be experimentally adjusted to obtain optimal values (see Section 4.4).

4.1 Objective Evaluation

For an objective evaluation, three normal dose CT images of abdomen, lung, and head are used as noise-free standard images, as shown in Figure 4(a,b,c). These testing images were cropped from original images of size 630×630 pixels of the same patient. Synthetic low-dose CT images are obtained by simulation by adding Gaussian noise with noise levels $\sigma = 10, 20$ and 30 to the standard images. The denoising methods perform denoising on these synthetic noisy low-dose images. Their results are then objectively compared to the original standard images based on the two well-known image quality assessment metrics, namely structural similarity (SSIM) [29] and peak signal-to-noise ratio (PSNR).

For MRFD, FD-SC1 and FD-SC2, for each of the testing CT images in Figure 4(a,b,c) we use three other standard images to construct standard patch databases (images in the three bottom rows in Figure 4). The patch size used in MRFD and FD-SC1 depends on the noise level, and thus in our experiments was set to $7 \times 7, 11 \times 11$ and 15×15 for the noise levels $\sigma = 10, 20$ and 30 , respectively. For FD-SC2, the patch size is fixed to 11×11 (the effect of the patch size is presented in Section 4.4). Parameter η in (13) is set to 0.

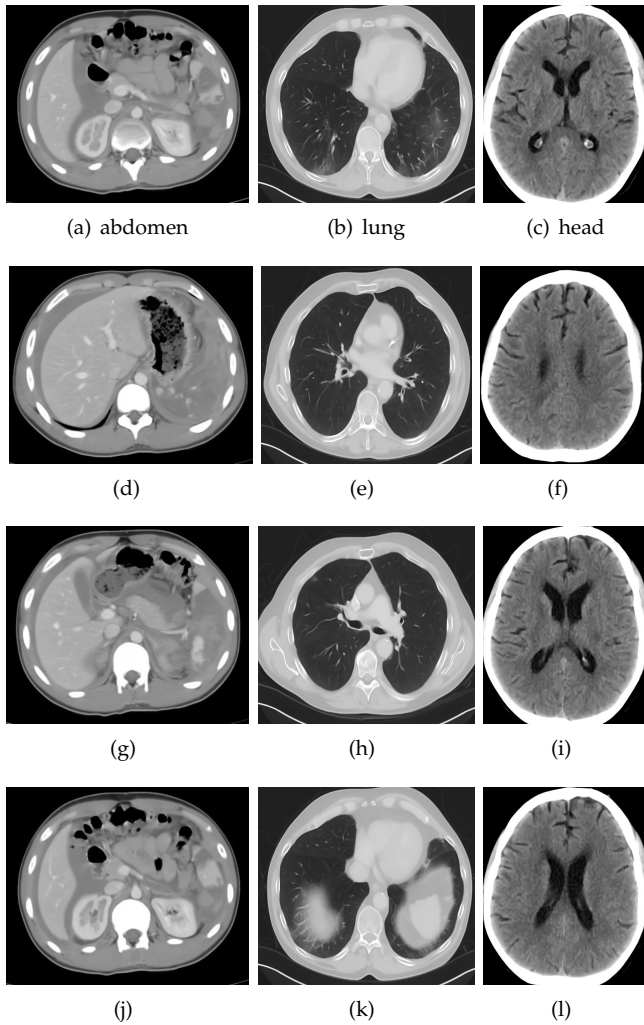


Figure 4. From left to right are CT images of abdomen, lung and head, respectively. The top rows contain testing images. These images are used to generate synthetic noisy low-dose images for objective evaluations. The remaining three rows contain the corresponding standard images. These images are used to build dictionaries for the learning-based denoising methods (MRFD, FD-SC1, FD-SC2). Source: <https://radiopaedia.org/cases>.

The best results of the methods are reported in Tables I and II in which the best values obtained for each noise level are in bold. As it can be seen, the quantitative evaluations show that FD-SC2 was almost superior to the other methods, indicating that our method is promising for denoising low-dose CT images.

4.2 Subjective Evaluation

For a subjective evaluation, we show in Figure 5 the experimental results on the CT image of the lung (Figure 4(b)) with noise level $\sigma = 20$. A region of interest (ROI) (the yellow rectangle) including a small point is zoomed in to facilitate visual comparison. This point is a subtle detail in the overall image. Globally, the methods denoised very effectively. However, it is observed that in the ROI in Figures 5(c)-5(e) the small point was smoothed out by TGV, NLM and KSVD. Although this small point was better preserved by MRFD, it is rather fuzzy (Figure 5(f)). Contrarily, as we can see in Figures 5(g) and 5(h), the small point in the ROI was effectively preserved by FD-SC1 and FD-SC2.

Table I
PSNR COMPARISON ON CT SCANS

Abdomen	TGV	NLM	KSVD	MRFD	FD-SC1	FD-SC2
$\sigma = 10$	34.2799	36.3421	36.8336	35.6138	36.4885	37.1032
$\sigma = 20$	30.8922	32.1371	31.9603	31.1545	31.8460	32.3604
$\sigma = 30$	28.2133	28.9184	28.6996	28.1523	28.6462	29.1911
Head	TGV	NLM	KSVD	MRFD	FD-SC1	FD-SC2
$\sigma = 10$	34.0854	34.5950	36.7655	35.4944	36.1263	36.9110
$\sigma = 20$	31.2148	31.6419	31.7591	31.3365	31.2083	32.0666
$\sigma = 30$	28.5300	28.9722	28.5374	28.3154	28.6859	29.1964
Lung	TGV	NLM	KSVD	MRFD	FD-SC1	FD-SC2
$\sigma = 10$	35.7034	37.3210	39.7222	36.7019	38.5793	38.8037
$\sigma = 20$	33.1212	34.4838	35.3297	32.9889	34.3846	35.6359
$\sigma = 30$	31.0489	31.6936	31.8397	30.1855	31.2497	32.9167

Table II
SSIM COMPARISON ON CT SCANS

Abdomen	TGV	NLM	KSVD	MRFD	FD-SC1	FD-SC2
$\sigma = 10$	0.7970	0.8083	0.7998	0.8089	0.8052	0.8213
$\sigma = 20$	0.7683	0.7141	0.7228	0.6929	0.7103	0.7436
$\sigma = 30$	0.65752	0.6426	0.6872	0.6078	0.6366	0.7018
Head	TGV	NLM	KSVD	MRFD	FD-SC1	FD-SC2
$\sigma = 10$	0.8347	0.8340	0.8714	0.8790	0.8817	0.8865
$\sigma = 20$	0.7895	0.7717	0.8045	0.7931	0.7781	0.8173
$\sigma = 30$	0.7464	0.7096	0.7668	0.7079	0.7393	0.7804
Lung	TGV	NLM	KSVD	MRFD	FD-SC1	FD-SC2
$\sigma = 10$	0.9496	0.9558	0.9695	0.9489	0.9619	0.9741
$\sigma = 20$	0.8877	0.8924	0.9439	0.8569	0.8923	0.9439
$\sigma = 30$	0.8669	0.8243	0.9134	0.7574	0.8038	0.9054

To further evaluate the effectiveness of the proposed method, we conducted experiments on the ELCAP public lung image database of the Cornell University¹. This database consists of real low-dose lung CT images and provides the locations of nodules detected by radiologists. As it can be seen in Figure 6(a), the low-dose image was strongly degraded by noise and artifacts.

In Figure 6, we show the denoising results obtained on one image in this database (Figure 6(a)). The low-dose CT image includes a small nodule (the position of nodules is zoomed and highlighted by rectangles). The patch database was constructed by using a normal-dose CT image (Figure 6(b)). Visually, noise in Figures 6(c)-6(d) was effectively denoised by TGV and NLM. However, the nodule was also almost suppressed. Compared to TGV and NLM, the nodule in Figure 6(f) obtained by MRFD was better preserved but slightly smoothed. For FD-SC1 and DF-SC2, we can easily see that they gave good results, as shown in Figures 6(g) and 6(h)). The small nodule was well preserved by FD-SC1 and FD-SC2. However, noise in Figure 6(h) (FD-SC2) seems to have been better removed, as compared to that in Figure 6(g) (FD-SC1). It is noticed that in this experiment we used only one normal-dose CT image (Figure 6(b)) to build the patch database for the learning-based methods (MRFD, FD-SC1 and FD-SC2). This demonstrates the potential of the proposed method.

¹<http://www.via.cornell.edu/lungdb.html>.

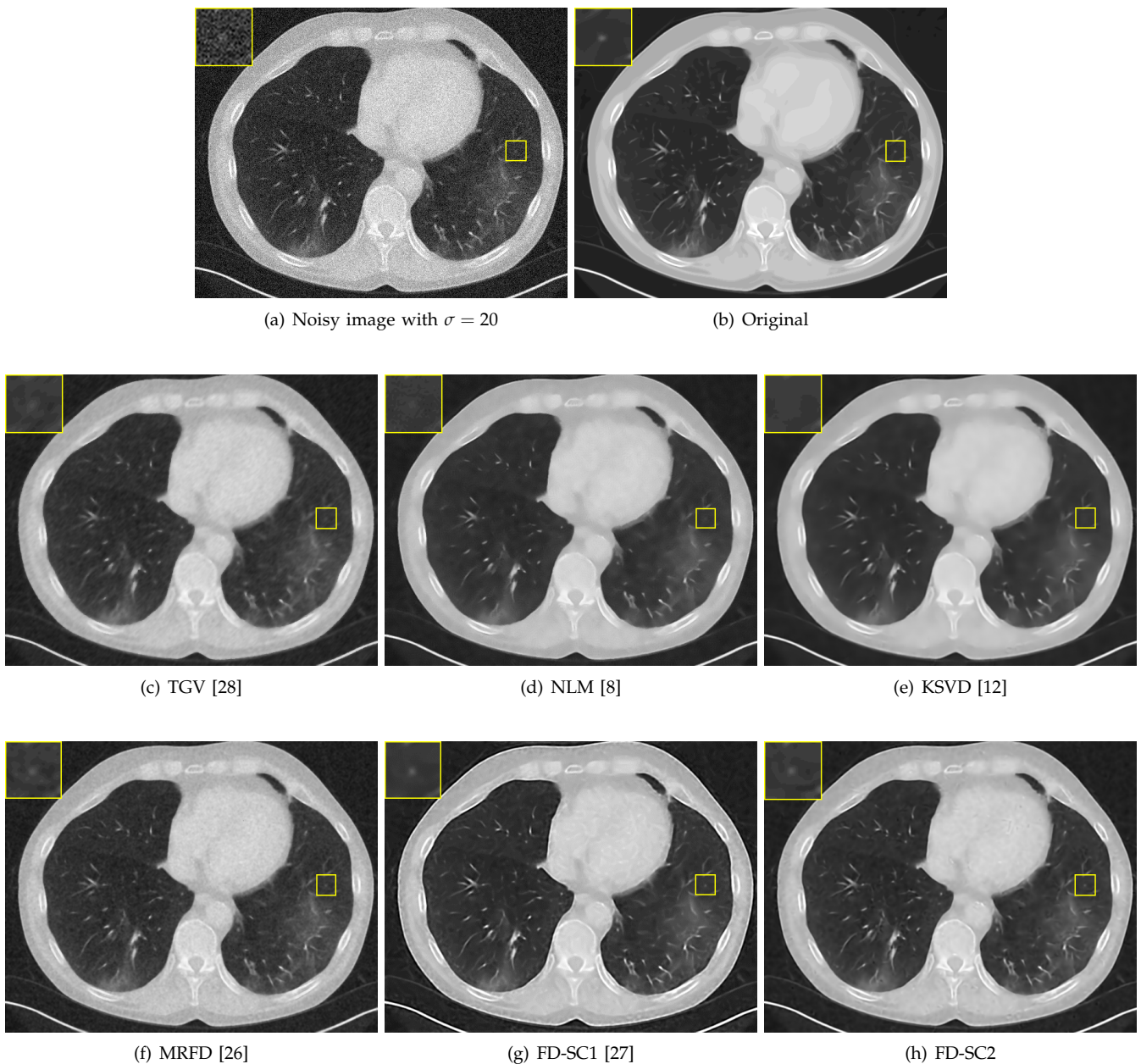


Figure 5. Results on synthetic low-dose CT image of lung with noise level $\sigma = 20$. A region of interest is zoomed in to visually compare, among different methods, the ability of preserving small details after denoising.

4.3 Computational Time

In order to compare computational complexity between the proposed method (FD-CS2) and its previous versions (MRFD and FD-SC1), we compute their running-time on the same image of the head as shown in Figure 4(c), with size of 608×481 for three different noise levels $\sigma = 10, 20$ and 30 . The patch databases were built from the same three standard images as shown in Figures 4(f), 4(i) and 4(l). The patch size was set to 9×9 for all noise levels $\sigma = 10, 20, 30$. The databases of the FD-CS1 and FD-CS2 included 972 patch-pairs. All tests were conducted using MATLAB 9.4 on a laptop with an Intel(R) Core(TM) i5 3210M CPU @ 2.50Gz, 4.0 GB Memory using Windows 64-bit.

In the denoising phase of FD-SC1 [27], for each noisy input patch, a set of K nearest neighbors was searched in the database before applying sparse coding using the sparse representation (2). This subset is considered as a sub-dictionary. In these experiments, the parameter K was set to 50 and the parameter L in (2) was set to 5.

Experimental results are reported in Table III. As it can be seen, FD-SC2 was significantly faster than MRFD and FD-SC1. Compared to MRFD [26], FD-SC2 was 13 times faster with $\sigma = 10$, 16.3 times faster with $\sigma = 20$, and 23 times faster with $\sigma = 30$. Compared to FD-SC1 [27], FD-SC2 was faster 4.8 times faster with $\sigma = 10$, 9.8 times faster with $\sigma = 20$, and 13.3 times faster with $\sigma = 30$.

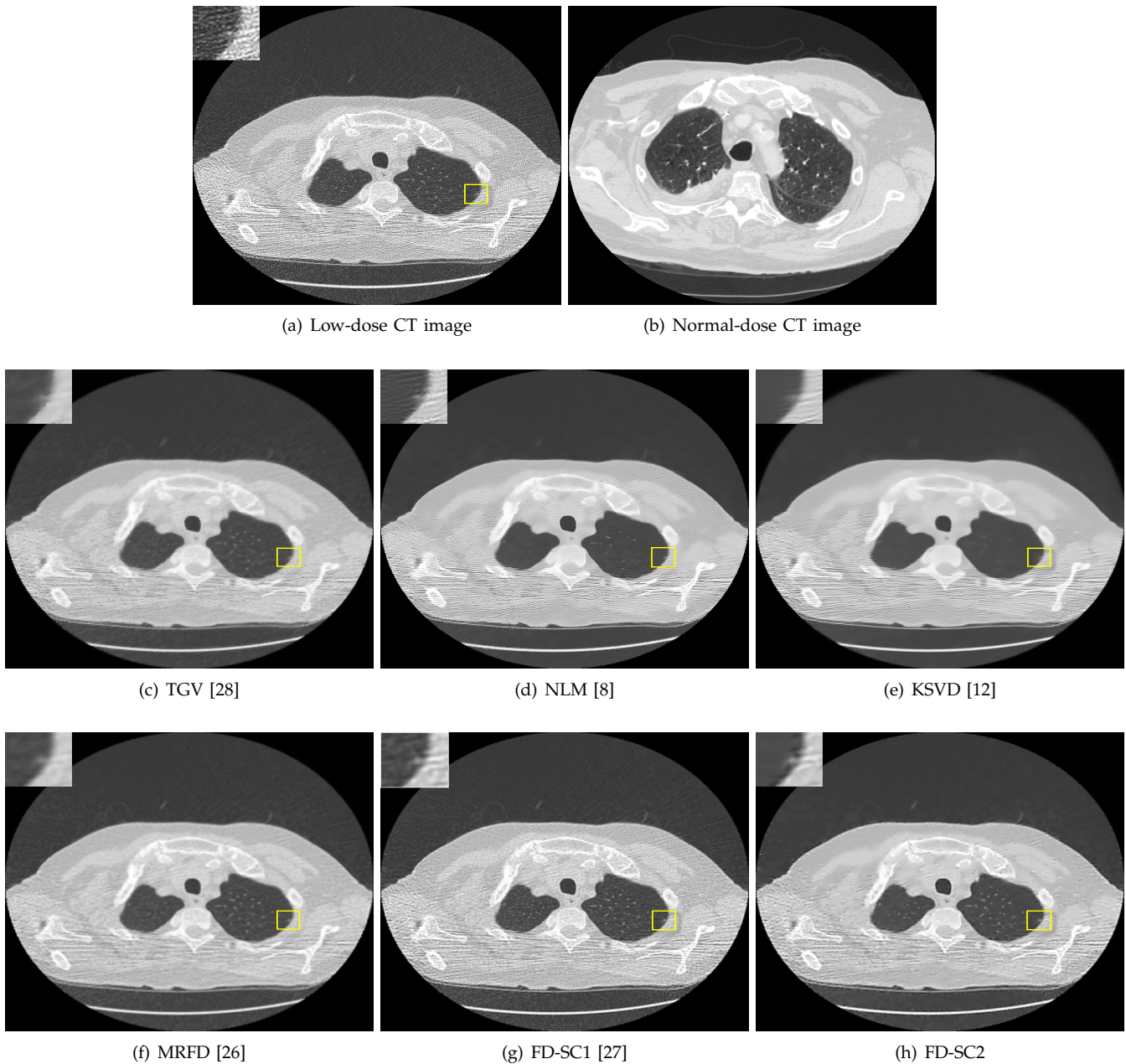


Figure 6. Denoising result on a real low-dose lung CT image including one nodule at position ($X=190, Y=394$).

Table III
RUNNING-TIME COMPARISON OF MRFD, FD-SC1 AND FD-SC2 WHEN
DENOISING THE HEAD IMAGE IN FIGURE 4(c) (SIZE 608×481) WITH
NOISE LEVELS $\sigma = 10, 20$ AND 30 (RUNNING-TIME IS GIVEN IN SECONDS)

Noise level	MRFD [26]	FD-SC1 [27]	FD-SC2
$\sigma = 10$	129.00	46.90	9.63
$\sigma = 20$	78.13	46.77	4.78
$\sigma = 30$	86.41	46.90	3.73

4.4 Empirical Study on Parameters

In this subsection, we present the effect of the threshold parameter ε in (11) and the patch size $\sqrt{n} \times \sqrt{n}$ on the performance of the proposed method (FD-SC2). Two well-known image quality metrics, SSIM

and PSNR, are used to analyze the influence of the parameters. Experiments were conducted on the three test images given in Figure 4 with three noise levels $\sigma = 10, 20$, and 30 .

To see the effect of γ , we performed experiments with $\gamma = 0.01, 0.02, \dots, 1.0$, and the patch size is fixed at 11×11 . The results, as shown in Figure 7, indicate that the optimal value of γ approximately was 0.09 for all three noise levels. It means that the optimal value of γ is less sensitive to noise.

Figure 8 shows the effect of the patch size. We performed experiments with different patch sizes, from 5×5 to 20×20 while keeping γ in these experiments constant at 0.09 . Visually, the optimal patch size in the the proposed method (FD-SC2) can be determined as 11×11 .

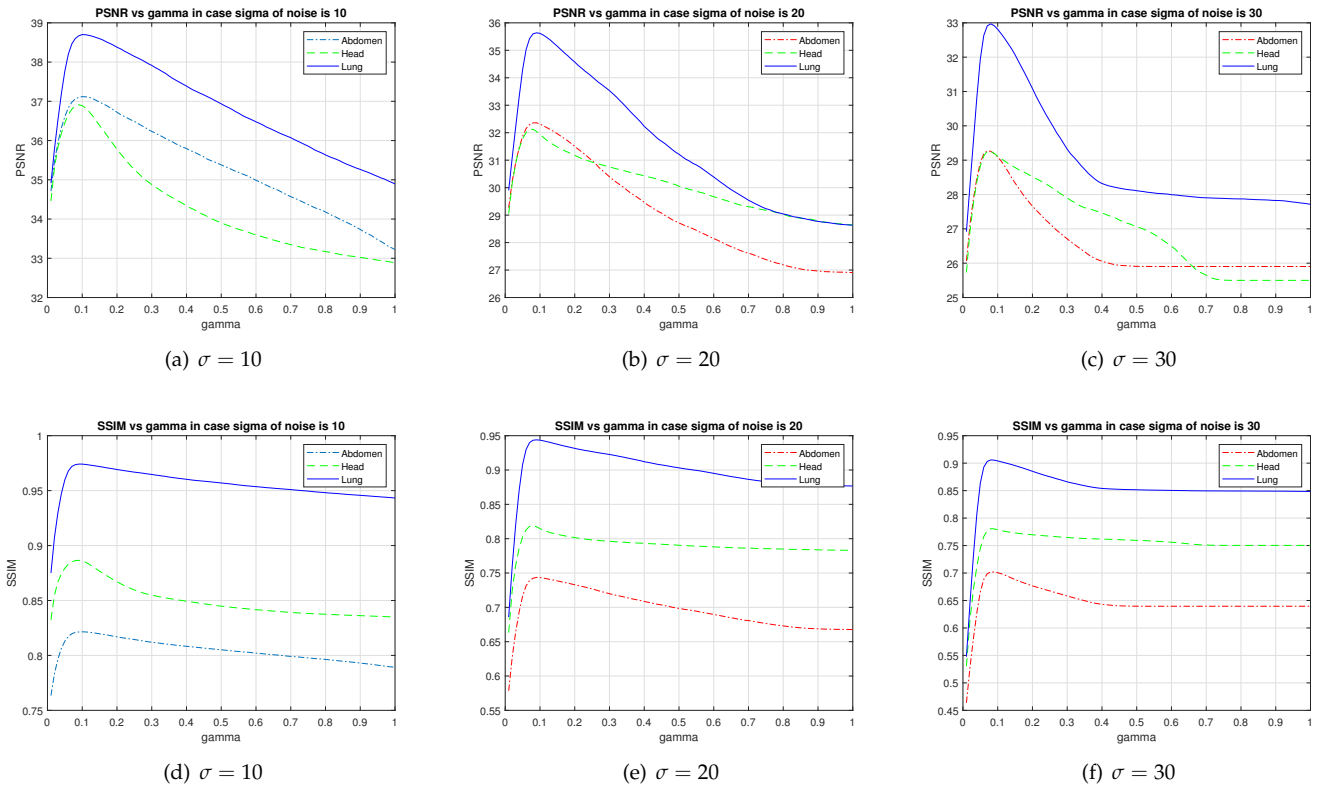


Figure 7. The effects of parameter γ in (11) on the PSNR and SSIM measures. Experiments were performed on images in Figure 4 with noise levels $\sigma = 10, 20, 30$.

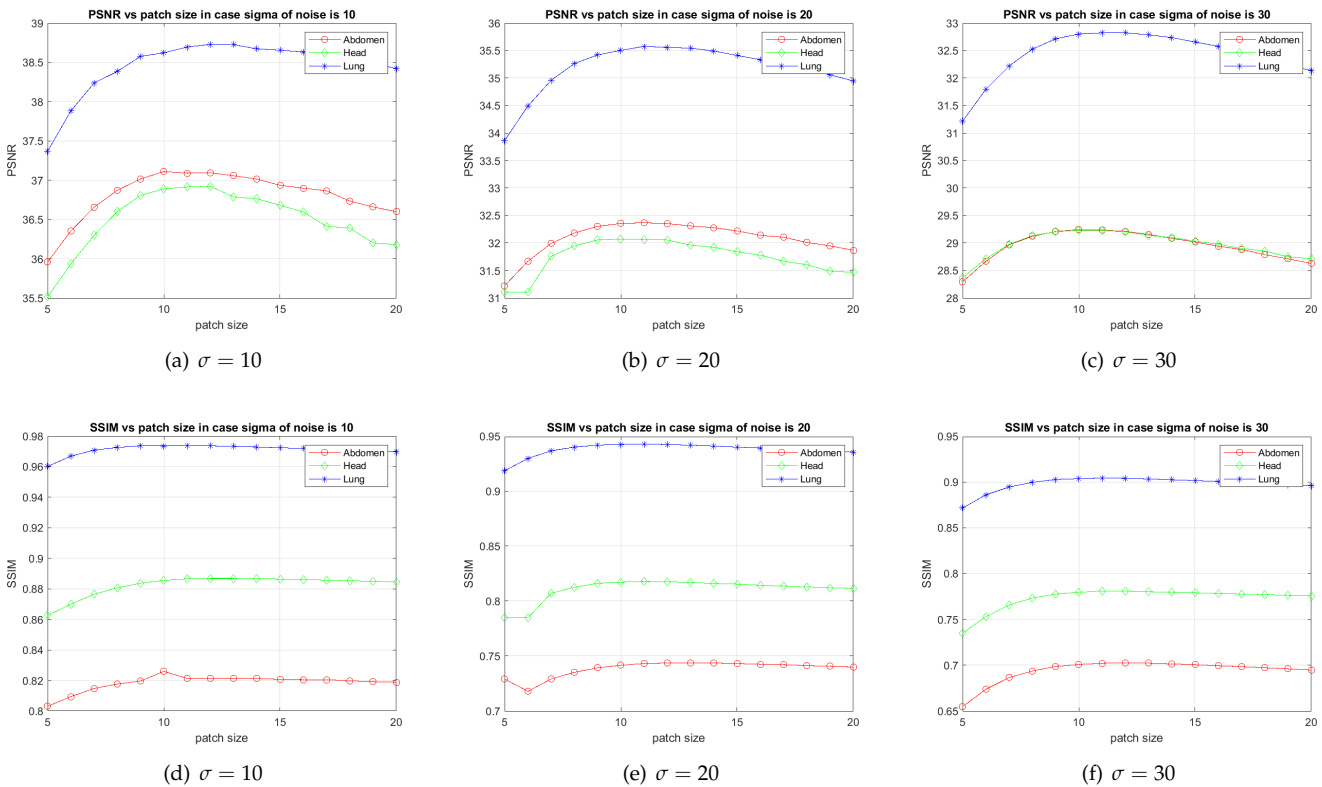


Figure 8. The effects of the patch size on the PSNR and SSIM measures. Experiments were performed on images in Figure 4 with noise levels $\sigma = 10, 20, 30$.

5 CONCLUSIONS

This paper has presented an effective denoising method for low-dose CT images. The proposed method, referred to as FD-SC2, was developed based on the assumption that noise mainly lies in the medium and high frequency components. Thus, a noisy image is decomposed into three component images of low, medium and high frequency bands, and then denoising is processed on the medium and high frequency component images. Moreover, inspired by recent success of learning-based methods, we have also exploited useful and available information from normal-dose CT images in FD-SC2. Useful information from the database of standard CT images is mined based on a sparse representation model. The effectiveness of the proposed method was verified through experimental results. FD-SC2 can effectively reduce noise while preserving well subtle image details as well as the nature of image. In future works, we will study the problem of patch dictionary optimization for sparse coding. Moreover, we will also seek a better solution for the image decomposition.

REFERENCES

- [1] G. Deng and L. Cahill, "An adaptive Gaussian filter for noise reduction and edge detection," in *Nuclear Science Symposium and Medical Imaging Conference, 1993 IEEE Conference Record*. IEEE, 1993, pp. 1615–1619.
- [2] F. Jin, P. Fieguth, L. Winger, and E. Jernigan, "Adaptive wiener filtering of noisy images and image sequences," in *Proceedings of the International Conference on Image Processing, ICIP 2003*, vol. 3. IEEE, 2003, pp. III–349.
- [3] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the Sixth International Conference on Computer Vision*. IEEE, 1998, pp. 839–846.
- [4] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [5] C. S. Drapaca, "A nonlinear total variation-based denoising method with two regularization parameters," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 582–586, 2009.
- [6] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.
- [7] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magnetic resonance in medicine*, vol. 65, pp. 480–491, 2011.
- [8] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2, 2005, pp. 60–65.
- [9] Z. Li, L. Yu, J. D. Trzasko, D. S. Lake, , D. J. Blezek, J. G. Fletcher, and C. H. McCollough, "Adaptive nonlocal means filtering based on local noise level for CT denoising," *Medical Physics*, vol. 41, no. 1, p. 011908, 2014.
- [10] M. Green, E. M. Marom, N. Kiryati, E. Konen, and A. Mayer, "Efficient low-dose CT denoising by locally-consistent non-local means (LC-NLM)," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016, Part III*, 2016, pp. 423–431.
- [11] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [12] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image processing*, vol. 15, no. 12, pp. 3736–3745, 2006.
- [13] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proceedings of the IEEE 12th International Conference on Computer Vision*. IEEE, 2009, pp. 2272–2279.
- [14] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1620–1630, 2013.
- [15] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [16] —, "BM3D image denoising with shape-adaptive principal component analysis," in *Proceedings of the SPARS'09-Signal Processing with Adaptive Sparse Structured Representations*, 2009.
- [17] D. Kang, P. J. Slomka, R. Nakazato, J. Woo, D. S. Berman, C. J. Kuo, and D. Dey, "Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm," in *Proceedings of the Medical Imaging 2013: Image Processing*, 2013, p. 86692G.
- [18] D. H. Trinh, M. Luong, J.-M. Rocchisani, C. D. Pham, and F. Dibos, "Medical image denoising using kernel ridge regression," in *Proceedings of the 18th IEEE International Conference on Image Processing (ICIP)*, 2011, pp. 1597–1600.
- [19] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2536–2545, 2017.
- [20] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.
- [21] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang, "Low dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Transactions on Medical Imaging*, 2018.
- [22] D. H. Trinh, M. Luong, J.-M. Rocchisani, C. D. Pham, H. D. Pham, and F. Dibos, "An optimal weight method for CT image denoising," *Journal of Electronic Science and Technology*, vol. 10, no. 2, pp. 124–129, 2012.
- [23] D.-H. Trinh, M. Luong, F. Dibos, J.-M. Rocchisani, C.-D. Pham, and T. Q. Nguyen, "Novel example-based method for super-resolution and denoising of medical images," *IEEE Transactions on Image processing*, vol. 23, no. 4, pp. 1882–1895, 2014.
- [24] Y. Chen, X. Yin, L. Shi, H. Shu, L. Luo, J.-L. Coatrieux, and C. Toumoulin, "Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing," *Physics in Medicine & Biology*, vol. 58, no. 16, p. 5803, 2013.
- [25] Y. Chen, L. Shi, Q. Feng, J. Yang, H. Shu, L. Luo, J. Coatrieux, and W. Chen, "Artifact suppressed dictionary learning for low-dose CT image processing," *IEEE Transactions on Medical Imaging*, vol. 33, no. 12, pp. 2271–2292, 2014.
- [26] D. H. Trinh, N. Linh-Trung, and T.-T. Nguyen, "An effective example-based denoising method for CT images using Markov random field," in *Proceedings of the International Conference on Advanced Technologies for Communications (ATC 2014)*, 2014, pp. 355–359.
- [27] T.-T. Nguyen, D.-H. Trinh, and N. Linh-Trung, "An efficient example-based method for CT image denoising based on frequency decomposition and sparse representation," in *International Conference on Advanced Technologies for Communications (ATC)*, 2016, pp. 293–296.
- [28] K. Bredies, K. Kunisch, and T. Pock, "Total generalized variation," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 492–526, 2010.

- [29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.



Nguyen Thanh Trung received his B.Eng. degree in Electronics and Telecommunications from the Faculty of Technology, Vietnam National University, Hanoi, Vietnam in 2003, M.S. degree in electronics engineering from University of Engineering and Technology, Vietnam National University, Hanoi, Vietnam in 2012. He is now a Phd student at VNU-UET and a lecturer at the Faculty of Electronics and Communication, University of Information and Communication Technology, Thain-guyen University, Vietnam. His research interests include biomedical signal and image processing, sparse coding, machine learning.



Trinh Dinh Hoan received the B.S. degree in mathematics education from the Vietnam National University, Hanoi, the M.S. degree in mathematics from the Hanoi Institute of Mathematics, and the Ph.D. degree in signal and image processing from the Laboratoire Analyse, Géométrie et Applications, Université Paris 13, Sorbonne Paris Cité, France, in 2005, 2008, and 2013, respectively. He is currently a Researcher with the Signals and Systems Laboratory, Université de Lorraine and CNRS, CRAN UMR 7039, 2 avenue de la Forêt de Haye, 54518 Vandœuvre-Lès-Nancy Cedex, France. His research interests include image restoration, super resolution, segmentation, sparse coding, machine learning, optimization, and medical imaging.



Nguyen Linh Trung obtained his B.Eng. and Ph.D. degrees, both in Electrical Engineering, from Queensland University of Technology, Brisbane, Australia. Since 2006, he has been on the faculty of University of Engineering and Technology, a member university of Vietnam National University, Hanoi, where he is currently an associate professor of electronic engineering in the Faculty of Electronics and Telecommunications and director of the Advanced Institute of Engineering and Technology. He is interested in signal processing methods, including time-frequency signal analysis, blind source separation, compressive sampling, tensor-based signal analysis, graph signal processing, and apply them to wireless communication and networking, biomedical engineering, with a current focus on large-scale processing.



Marie Luong received the Engineer Diploma degree from the Ecole Nationale Supérieure d'Electricité et de Mécanique de Nancy, France, and the Ph.D. degree in Automatic, Diagnosis, Signal processing from the L'Institut National Polytechnique de Lorraine (Lorraine University), France, in 1991 and 1996, respectively. She is currently an Associate Professor with the Université Paris 13, Sorbonne Paris Cité, France. Her research interests include image restoration, superresolution, segmentation, watermarking, and medical imaging.