Fractional-order Sparse Representation for Image Denoising

Leilei Geng, Zexuan Ji, Yunhao Yuan, and Yilong Yin

***Abstract*—Sparse representation models have been shown promising results for image denoising. However, conventional sparse representation-based models cannot obtain satisfactory estimations for sparse coefficients and the dictionary. To address this weakness, in this paper, we propose a novel fractional-order sparse representation (FSR) model. Specifically, we cluster the image patches into K groups, and calculate the singular values for each clean/noisy patch pair in the wavelet domain. Then the uniform fractional-order parameters are learned for each cluster. Then a novel fractional-order sample space is constructed using adaptive fractional-order parameters in the wavelet domain to obtain more accurate sparse coefficients and dictionary for image denoising. Extensive experimental results show that the proposed model outperforms state-of-the-art sparse representation-based models and the block-matching and 3D filtering algorithm in terms of denoising performance and the computational efficiency.**

***Index Terms*—Fractional-order, image denoising, multi-scale, sparse representation.**

I. Introduction

IMAGE denoising is one of the most important research subjects in image processing due to its fundamental role in a number of applications, such as medical imaging, remote sensing, surveillance and entertainment. Generally, the observed noisy image $\boldsymbol{y}$ can be formulated as $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{x}$ is the clean image and $\boldsymbol{n}$ is an additive white Gaussian noise with zero-mean and standard deviation σ. In the past decades, extensive studies have been conducted to develop various image denoising approaches, such as total variation (TV) regularization [1], non-local means (NLM) algorithm [2] and block-matching and 3D (BM3D) filtering algorithm [3]. Although classical approaches are effective in removing noise artifacts, the TV regularization tend to over-smooth images due to the piecewise constant assumption and the computational complexity of the BM3D is extremely high. As an alternative, sparse representation-based models have recently led to promising results [4]−[7].

Sparse representation-based models assume that an image $\boldsymbol{x}$ can be represented as $\boldsymbol{x} \approx D\boldsymbol{\alpha}$, where $D \in \mathbb{R}^{n\times m}$ is an over-complete dictionary and most entries of the sparse coefficients α are zero or close to zero. The sparse decomposition of $\boldsymbol{x}$ can be obtained by solving an l_0-minimization problem, i.e., $\boldsymbol{\alpha} = \arg\min_\alpha \|\boldsymbol{\alpha}\|_0$, s.t. $\|\boldsymbol{x} - D\boldsymbol{\alpha}\|_2 \leq \varepsilon$, where $\|\cdot\|_0$ is a pseudo norm and ε is a small constant. Since l_0-minimization is a NP (none-deterministic polynomial)-hard combinatorial optimization problem, it is often relaxed to l_1-minimization. The l_1-norm based model can be solved efficiently by convex optimization techniques, formulated as $\{\boldsymbol{\alpha}, D\} = \arg\min_\alpha\{\|\boldsymbol{x} - D\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1\}$, where λ denotes the regularization parameter. Many efficient l_1-minimization techniques have been proposed, such as iterative thresholding algorithms [8], [9], Bregman split algorithms [10], [11] and so on.

Dictionary learning is a key problem of sparse representation-based models. Comparing with analytically designed dictionaries (e.g., the wavelet/curvelet dictionary), dictionaries learned from example image can greatly improve the denoising performance because they can better characterize the image structures [12]−[16]. For example, [16] estimated the sparse coefficients by exploiting nonlocal self-similarity in the observed image. Recently, most conventional sparse representation-based models concentrate on learning the adaptive dictionary directly from the noisy image itself instead of from the external image database. The main advantages are that the self-structural information is exploited and the trained dictionary is adaptive to the image of interest. However, the training samples extracted directly from the observed image are inevitably corrupted by noise. Thus, accurate sparse coefficients and optimal dictionary cannot be obtained efficiently in the noisy sample space. Consequently, the performance and efficiency of conventional sparse representation-based models are limited.

In this paper, we propose a fractional-order sparse representation (FSR) model to obtain better estimations for both the sparse coefficients and the dictionary. First, we cluster external training image patches into multiple groups to explore image self-similarity. Then we calculate the singular values for each clean/noisy patch pair in the wavelet domain and learn the uniform fractional-order parameters for each cluster. To deal with

Manuscript received March 1, 2016; accepted October 26, 2016. This work was supported by the National Natural Science Foundation of China (61573219, 61402203, 61401209, 61701192, 61671274), the Opening Fund of Shandong Provincial Key Laboratory of Network Based Intelligent Computing, the Fostering Project of Dominant Discipline and Talent Team of Shandong Province Higher Education Institutions, and Fostering Project of Dominant Discipline and Talent Team of SDUFE. Recommended by Associate Editor Minling Zhang. *(Corresponding author: Leilei Geng.)*

Citation: L. L. Geng, Z. X. Ji, Y. H. Yuan, and Y. L. Yin, "Fractional-order sparse representation for image denoising," *IEEE/CAA J. of Autom. Sinica*, vol. 5, no. 2, pp. 555−563, Mar. 2018.

L. L. Geng and Y. L. Yin are with the School of Computer Science and Technology, Shandong University of Finance and Economics, Jinan 250014, China (e-mail: leileigeng_njust@163.com; ylyin@sdu.edu.cn).

Z. X. Ji is with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: jizexuan@njust.edu.cn).

Y. H. Yuan is with the School of Internet of Things Engineering, Jiangnan University, Wuxi 214112, China (e-mail: yyhzbh@163.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2017.7510412

the noisy image, a novel sample space is constructed using adaptive fractional-order parameters in the wavelet domain. In this way, we can obtain more accurate sparse coefficients and dictionary for image denoising. Comparing with several state-of-the-art sparse representation-based models [4], [16] and the BM3D algorithm [3], our proposed FSR model clearly performs better in terms of both denoising performance and computational efficiency.

II. Background and Related Work

A. Sparse Representation-based Denoising Model

Given the dictionary $D_{\boldsymbol{x}}$, an image patch $\boldsymbol{x}_i = R_i\boldsymbol{x}$, extracted from the clean image $\boldsymbol{x}$ at the location i, can be sparsely represented by solving the non-convex l_1-optimization:

$$\{\boldsymbol{\alpha}_{\boldsymbol{x},i}, D_{\boldsymbol{x}}\} = \arg\min_{\alpha_i}\{\|\boldsymbol{x}_i - D\boldsymbol{\alpha}_i\|_2^2 + \lambda\|\boldsymbol{\alpha}_i\|_1\}. \quad (1)$$

Then, the target image $\boldsymbol{x}$ can be constructed by averaging each image patch $\boldsymbol{x}_i$ with a straightforward least-square solution in [12]:

$$\widehat{\boldsymbol{x}} \approx D_{\boldsymbol{x}} \circ \boldsymbol{\alpha}_{\boldsymbol{x}} = (\sum_i R_i^T R_i)^{-1} \sum_i (R_i^T D_{\boldsymbol{x}} \boldsymbol{\alpha}_{\boldsymbol{x},i}). \quad (2)$$

The conventional sparse representation (SR) denoising models recover $\boldsymbol{x}$ from noisy image $\boldsymbol{y}$ by solving the following minimization problem:

$$\{\boldsymbol{\alpha}_{\boldsymbol{y}}, D_{\boldsymbol{y}}\} = \arg\min_{\alpha}\{\|\boldsymbol{y} - D\boldsymbol{\alpha}\|_2^2 + \lambda\|\boldsymbol{\alpha}\|_1\}. \quad (3)$$

Then, by adding auxiliary variable $\boldsymbol{x}_i = D_{\boldsymbol{y}}\boldsymbol{\alpha}_{\boldsymbol{y},i}$, the denoised image $\widehat{\boldsymbol{x}}$ is reconstructed with a closed-form solution in [17]:

$$\widehat{\boldsymbol{x}} \approx D_{\boldsymbol{y}} \circ \boldsymbol{\alpha}_{\boldsymbol{y}} = (\lambda I + \sum_i R_i^T R_i)^{-1} \sum_i (\lambda \boldsymbol{y} + R_i^T D_{\boldsymbol{y}} \boldsymbol{\alpha}_{\boldsymbol{y},i}). \quad (4)$$

B. Fractional-order Technique

In recent years, the fractional-order technique has been widely used in image denoising [18], texture enhancement [19], image reconstruction [20], face representation [21], and other applications. Reference [18] proposed a fractional-varying-order differential-based method for image denoising. This method varied the gradient of images to suppress the blocky effect. Moreover, it retained more texture details than the integral order partial differential denoising model due to the fractional differential amplitude frequency. Reference [19] presented a fractional differential-based method for multi-scale texture enhancement. This method could nonlinearly enhance complex texture details, which has been proven superior to traditional integral-based algorithms. Reference [20] incorporated fractional derivatives and edge detection in a method called fractional-order partial differential-based image reconstruction. The fractional derivative has been demonstrated to be superior to the integer derivative. In face representation, [21] proposed a fractional-order embedding canonical correlation analysis to reduce the dimensionality of multi-view data. This method constructed the fractional-order both within-set and between-set scatter matrices to alleviate the deviation of the eigenvalues and singular values in the corresponding sample covariance matrices. We can conclude that the fractional-order technique targets at a more general framework of nonlinear and flexible data processing mode. Therefore, we introduce the fractional-order technique into the sparse representation model to obtain more accurate sparse coefficients and dictionary.

III. Modeling of Fractional-order Sparse Representation

A. Motivation

For effective image denoising, the dictionary $D_{\boldsymbol{y}}$ and the sparse coefficients $\boldsymbol{\alpha}_{\boldsymbol{y}}$ obtained by solving the minimization problem in (3) are respectively expected to be as close as possible to the $D_{\boldsymbol{x}}$ and the $\boldsymbol{\alpha}_{\boldsymbol{x}}$ obtained by (1). However, due to the noise in the observed image, $D_{\boldsymbol{y}}$ and $\boldsymbol{\alpha}_{\boldsymbol{y}}$ will deviate from their true values. Thus, both $D_{\boldsymbol{y}}$ and $\boldsymbol{\alpha}_{\boldsymbol{y}}$ cannot be considered as good approximations, and would directly decrease the quality of the denoised image. To further illustrate these difference, we define the deviation degree of the dictionary $D_{\boldsymbol{y}}$ as $\mathrm{Dev}(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ and the deviation degree of the sparse coefficients $\boldsymbol{\alpha}_{\boldsymbol{y}}$ as $\mathrm{Dev}(\boldsymbol{\alpha}_{\boldsymbol{x}}, \boldsymbol{\alpha}_{\boldsymbol{y}})$.

Definition 1: Suppose $D_{\boldsymbol{x}} = \{\boldsymbol{d}_{\boldsymbol{x},1}, \boldsymbol{d}_{\boldsymbol{x},2}, \ldots, \boldsymbol{d}_{\boldsymbol{x},m}\}$ is an over-complete dictionary containing m dictionary atoms learned from the original image $\boldsymbol{x}$, and $D_{\boldsymbol{y}} = \{\boldsymbol{d}_{\boldsymbol{y},1}, \boldsymbol{d}_{\boldsymbol{y},2}, \ldots, \boldsymbol{d}_{\boldsymbol{y},m}\}$ is an over-complete dictionary containing m dictionary atoms learned from the noisy image $\boldsymbol{y}$. Then, the deviation degree of the dictionary $D_{\boldsymbol{y}}$ is defined as:

$$\mathrm{Dev}(D_{\boldsymbol{x}}, D_{\boldsymbol{y}}) = \{\boldsymbol{dev}_i\}_{i=1}^m = \left\{1 - \max_{1\le j\le m} \frac{|\boldsymbol{d}_{\boldsymbol{x},j} \cdot \boldsymbol{d}_{\boldsymbol{y},i}|}{|\boldsymbol{d}_{\boldsymbol{x},j}| \cdot |\boldsymbol{d}_{\boldsymbol{y},i}|}\right\}_{i=1}^m \quad (5)$$

where $\mathrm{Dev}(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ denotes the deviation degree vector of the dictionary $D_{\boldsymbol{y}}$, $\boldsymbol{dev}_i$ denotes the ith deviated element referring to the ith dictionary atom $\boldsymbol{d}_{\boldsymbol{y},i}$, and $\boldsymbol{d}_{\boldsymbol{x},j} \cdot \boldsymbol{d}_{\boldsymbol{y},i}$ denotes the inner product of $\boldsymbol{d}_{\boldsymbol{x},j}$ and $\boldsymbol{d}_{\boldsymbol{y},i}$. The definition of $\mathrm{Dev}(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ describes the ability that the dictionary $D_{\boldsymbol{y}}$ express the noisy image $\boldsymbol{y}$. When the $\mathrm{Dev}(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ is zero, the $D_{\boldsymbol{y}}$ is as the same as the $D_{\boldsymbol{x}}$ and the noise can be restrained effectively.

Definition 2: Suppose that $\boldsymbol{\alpha}_{\boldsymbol{x}}$ is the sparse coefficients of the clean image patch and $\boldsymbol{\alpha}_{\boldsymbol{y}}$ is the sparse coefficients of the noise image patch. Then, the deviation degree of the sparse coefficients $\boldsymbol{\alpha}_{\boldsymbol{y}}$ is defined as:

$$\mathrm{Dev}(\boldsymbol{\alpha}_{\boldsymbol{x}}, \boldsymbol{\alpha}_{\boldsymbol{y}}) = \frac{\|\boldsymbol{\alpha}_{\boldsymbol{y}} - \boldsymbol{\alpha}_{\boldsymbol{x}}\|_F}{\|\boldsymbol{\alpha}_{\boldsymbol{x}}\|_F} \quad (6)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. The definition of $\mathrm{Dev}(\boldsymbol{\alpha}_{\boldsymbol{x}}, \boldsymbol{\alpha}_{\boldsymbol{y}})$ describes the similarity between $\boldsymbol{\alpha}_{\boldsymbol{x}}$ and $\boldsymbol{\alpha}_{\boldsymbol{y}}$. As the $\mathrm{Dev}(\boldsymbol{\alpha}_{\boldsymbol{x}}, \boldsymbol{\alpha}_{\boldsymbol{y}})$ is close to zero, the sparse coefficients $\boldsymbol{\alpha}_{\boldsymbol{y}}$ is similar to its corresponding true value $\boldsymbol{\alpha}_{\boldsymbol{x}}$.

To show the deviation degrees, we conducted experiments using the Lena image. We added Gaussian white noise (σ =20) to the original image $\boldsymbol{x}$ to obtain the noisy image $\boldsymbol{y}$. For fair comparison, both $\boldsymbol{\alpha}_{\boldsymbol{x}}$ and $\boldsymbol{\alpha}_{\boldsymbol{y}}$ were obtained using fixed discrete cosine transform (DCT) bases. Figs. 1(a) and 1(b) plot the distributions of $\mathrm{Dev}(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ and $\mathrm{Dev}(\boldsymbol{\alpha}_{\boldsymbol{x}}, \boldsymbol{\alpha}_{\boldsymbol{y}})$, respectively.

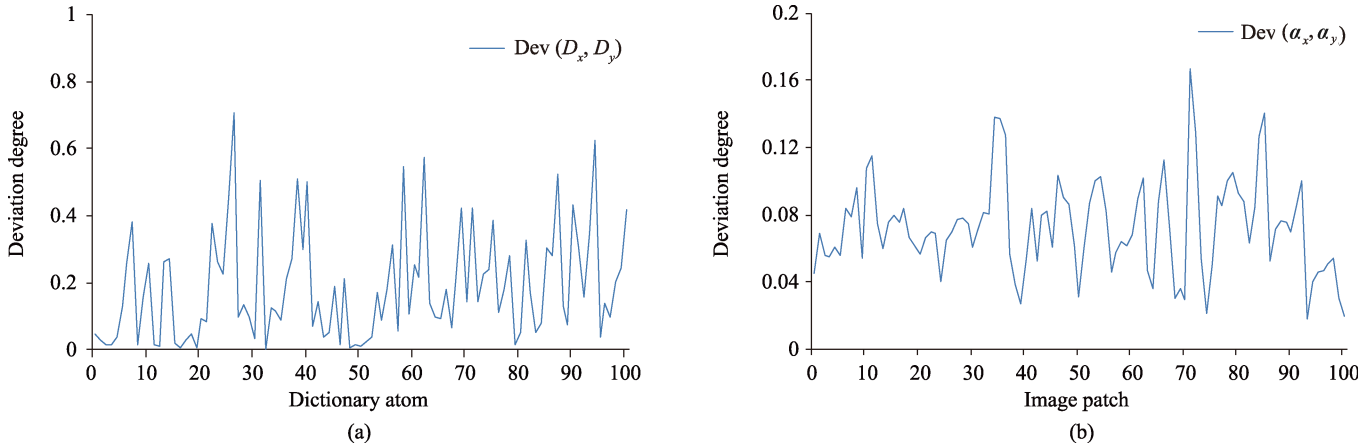


Fig. 1. Distributions of Dev$(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ and Dev$(\boldsymbol{\alpha_x}, \boldsymbol{\alpha_y})$ when the Lena image is noisy ($\sigma = 20$).

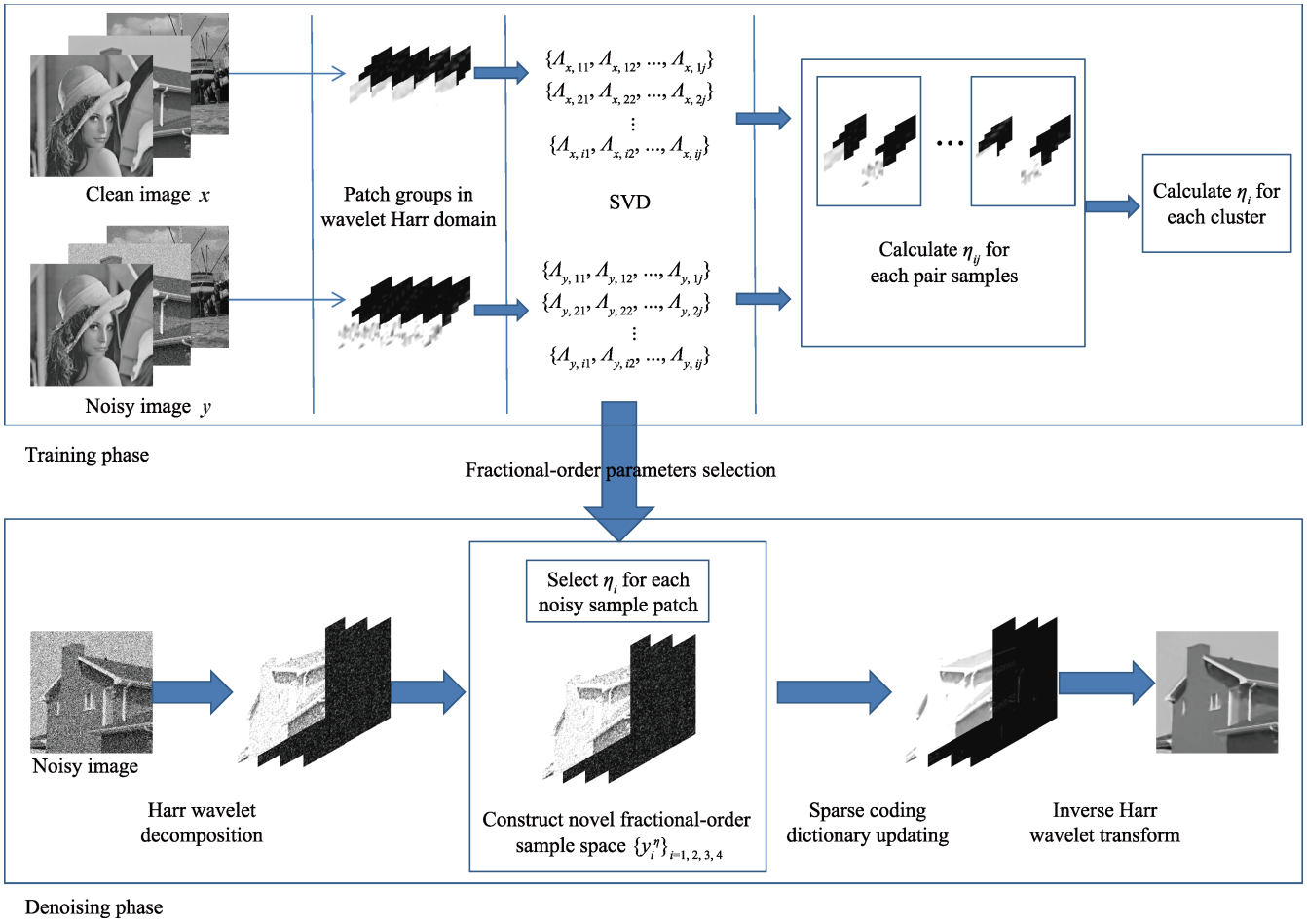


Fig. 2. Flowchart of the proposed FSR model.

Note that most deviation degrees of dictionary atoms are over 0.1 and the highest value is about 0.71. As shown in Fig. 1(b), more than 80 % of Dev$(\boldsymbol{\alpha_x}, \boldsymbol{\alpha_y})$ concentrated in 0.05−0.12. Hence the accurate $\boldsymbol{\alpha_y}$ and optimal $D_{\boldsymbol{y}}$ cannot be obtained from the noisy image by the SR models. The definitions of Dev$(D_{\boldsymbol{x}}, D_{\boldsymbol{y}})$ and Dev$(\boldsymbol{\alpha_x}, \boldsymbol{\alpha_y})$ indicate that it is possible to further improve the denoising performance by improving the accuracy of both $\boldsymbol{\alpha_y}$ and $D_{\boldsymbol{y}}$. However, it is difficult to directly improve the accuracy of $\boldsymbol{\alpha_y}$ and $D_{\boldsymbol{y}}$ in the observed sample space due to noise disturbance. To address this problem, we construct a novel sample space by correcting singular values utilizing the cluster-based fractional-order technique with adaptive parameters in the wavelet domain.

B. Fractional-order Sparse Representation Model

The FSR model involves the training phase and the denoising phase:

In the training phase, we learn the fractional-order parameters from external image database including four steps:

Step 1: We cluster the training image patches into K clusters (typically K=10) using K-means algorithm and decompose those patches into frequency sub-patches by the discrete Haar wavelet transform;

Step 2: We calculate singular values for each patch;

Step 3: The fractional-order parameter $\boldsymbol{\eta}_{ij}$ of each patch is obtained given the clean/noisy patch pairs;

Step 4: We compute the uniform fractional-order parameter $\boldsymbol{\eta}_i$ for each training patch group with similar textures.

In the denoising phase, the noisy image is first decomposed into Haar wavelet images. Then the fractional-order parameter is employed to construct the novel fractional-order sample space, and the denoised image is reconstructed by sparse coefficients and dictionary. The flowchart of the FSR model is shown in Fig. 2.

1) Deviation Degree of Singular Values: For an image patch $\boldsymbol{x}_i$, we have the singular value decomposition (SVD) [22] as $\boldsymbol{x}_i = P_i \Lambda_i Q_i$, where P_i and Q_i are the left and right singular matrices, respectively, $\Lambda_i = \mathrm{diag}\{s_{i,1}, s_{i,2}, \ldots, s_{i,\sqrt{n}}\}$ is a singular value matrix, and $s_{i,1} \geq s_{i,2} \geq \cdots \geq s_{i,\sqrt{n}}$ are the descending order singular values. The singular values can be interpreted as the variance estimation and they expose the geometric structure of the image. Hence, the singular values represent the intrinsic and algebraic properties containing illumination, noise and texture. Due to the disturbance of noise, the singular values obtained from the noisy image deviate from their true values. Similarly, we can define the deviation degree of singular values as $\mathrm{Dev}(\Lambda_{\boldsymbol{x}}, \Lambda_{\boldsymbol{y}})$.

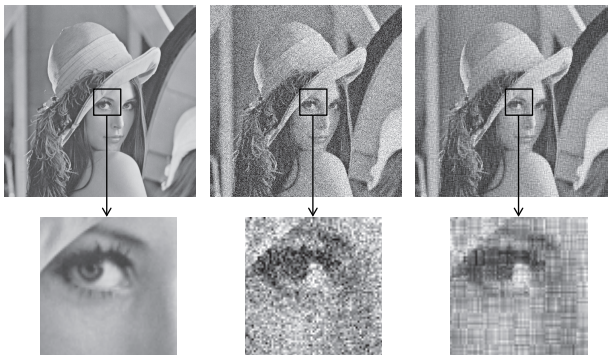
Definition 3: Suppose $\Lambda_{\boldsymbol{x}} = \mathrm{diag}\{s_{\boldsymbol{x},1}, s_{\boldsymbol{x},2}, \ldots, s_{\boldsymbol{x},\sqrt{n}}\}$ denotes the singular value matrix of the original image patch, and $\Lambda_{\boldsymbol{y}} = \mathrm{diag}\{s_{\boldsymbol{y},1}, s_{\boldsymbol{y},2}, \ldots, s_{\boldsymbol{y},\sqrt{n}}\}$ denotes the singular value matrix of the noise image patch. Then, the deviation degree of the singular values is defined as:

$$\mathrm{Dev}(\Lambda_{\boldsymbol{x}}, \Lambda_{\boldsymbol{y}}) = \frac{\|\Lambda_{\boldsymbol{y}} - \Lambda_{\boldsymbol{x}}\|_F}{\|\Lambda_{\boldsymbol{x}}\|_F}. \tag{7}$$

Fig. 3. Deviation degree of singular values.

To show the deviation degree of the singular values, we plotted the distribution of $\mathrm{Dev}(\Lambda_{\boldsymbol{x}}, \Lambda_{\boldsymbol{y}})$ in Fig. 3, where $\Lambda_{\boldsymbol{y}}$ was computed from the noisy Lena image ($\sigma = 20$). As observed in Fig. 3, it is evident that the singular values differ greatly from their true values and most $\mathrm{Dev}(\Lambda_{\boldsymbol{x}}, \Lambda_{\boldsymbol{y}})$ are over 4. Recently, this deviation has been demonstrated to have a negative effect on classification systems [21]. Motivated by the above situation, we re-estimate the singular values using fractional-order techniques.

2) Multi-scale Decomposition: As mentioned above, due to disturbance of noise, the singular values deviate from their true values. The denoised image would involve unexpected perturbations when the singular values are directly obtained from the noisy image space. As illustrated in Fig. 4, the re-estimated image Fig. 4(c) is directly reconstructed using the corresponding singular values obtained from the original Lena image Fig. 4(a). Although the extrinsic disturbance (noise) is restrained (e.g., Fig. 4(b)), other new disturbances are involved. Because the extrinsic factor (noise) can be largely isolated from the intrinsic factor (texture) in the frequency sub-spaces, the re-estimated sample space can avoid these unexpected perturbations. Thus, it is necessary to correct the singular values of noisy image patches in the wavelet domain.



(a) Original image (b) Noisy image (c) Reconstructed image

Fig. 4. Comparison of experimental results.

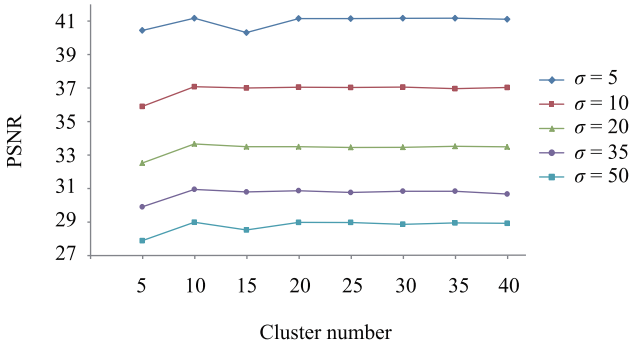


Fig. 5. Comparison of PSNRs with various cluster numbers and various noise levels.

3) Fractional-order Parameter Learning: Since the proposed FSR model involves the fractional-order technique, the key problem is that how to choose the fractional-order parameter $\boldsymbol{\eta}$. In this paper, we first cluster the training image patches into K groups using the K-means clustering method. For each patch in the ith cluster, we calculate the fractional-order by $\boldsymbol{\eta}_{ij} = \arg\min \mathrm{Dev}(\Lambda_{\boldsymbol{x}}, \Lambda_{\boldsymbol{y}})$. Then the uniform fractional-order parameter $\boldsymbol{\eta}_i$ for each training group is obtained by $\boldsymbol{\eta}_i = \arg\min_{\boldsymbol{\eta}_i} \sum \|\boldsymbol{\eta}_i - \boldsymbol{\eta}_{ij}\|_F^2$. Furthermore, the cluster number K (typically K=10) is determined by a series of experiments. We utilize the cross-validation method to determine the parameter K with optimal peak signal to noise ratio (PSNR) of the denoised image. Fig. 5 plots the distribution of PSNR average results over five typical images

(e.g., Lena/Barbara/Boat/House/Pepper). We can see that there is little improvement of the image denoising performance over all noise levels as the cluster number increases. Therefore, we cluster the samples into 10 groups in terms of the effectiveness and efficiency.

4) Fractional-order Image Denoising: With the trained fractional-order parameters, we denoise the image in the novel fractional-order sample space. Now, we define the fractional-order patch as follows:

Definition 4: Suppose $\boldsymbol{\eta} = \{\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \ldots, \boldsymbol{\eta}_{\sqrt{n}}\}$ is a fractional-order parameter, and each fractional element $\boldsymbol{\eta}_k$ satisfies $0 \leq \boldsymbol{\eta}_k \leq 1$. Then, the fractional-order patch $\boldsymbol{y}_i^{\boldsymbol{\eta}}$ is referred to as:

$$\boldsymbol{y}_i^{\boldsymbol{\eta}} = P_i \Lambda_i^{\boldsymbol{\eta}} Q_i,\ \Lambda_i^{\boldsymbol{\eta}} = \mathrm{diag}(s_{i,1}^{\boldsymbol{\eta}_1}, s_{i,2}^{\boldsymbol{\eta}_2}, \ldots, s_{i,\sqrt{n}}^{\boldsymbol{\eta}_{\sqrt{n}}}) \tag{8}$$

where P_i and Q_i are respectively the left and right singular matrices of the noisy patch $\boldsymbol{y}_i$, and $\Lambda_i^{\boldsymbol{\eta}}$ denotes the fractional-order singular value matrix. In Definition 4, it is obvious that the decomposition in (8) subsume the thin SVD as a special case when $\boldsymbol{\eta} = \{1, 1, \ldots, 1\}$. Based on Definition 4, the following properties are observed:

Property 1: The rank of $\boldsymbol{y}_i^{\boldsymbol{\eta}}$ is equal to the rank of $\boldsymbol{y}_i$, i.e., rank $(\boldsymbol{y}_i^{\boldsymbol{\eta}})$ = rank$(\boldsymbol{y}_i)$.

Property 2: $\|\boldsymbol{y}_i^{\boldsymbol{\eta}} - \boldsymbol{y}_i\|_F^2 = \sum_{m=1}^{\sqrt{n}} (s_m^{\boldsymbol{\eta}_m} - s_m)^2$, where $\|\cdot\|_F$ denotes the Frobenius norm.

Property 1 shows the basic characteristics of the novel fractional-order sample. Property 2 reveals that the space distance between $\boldsymbol{y}_i^{\boldsymbol{\eta}}$ and $\boldsymbol{y}_i$ can be controlled by adjusting the fractional-order parameter $\boldsymbol{\eta}$.

After constructing fractional-order patches, we build a novel sparse representation-based model for image denoising. The proposed fractional-order image denoising phase is an iterative algorithm that alternates between fractional-order sample space construction and dictionary learning. In the fractional-order sample space construction stage, the noisy image is first decomposed into different frequency sub-spaces by the discrete Haar wavelet transform with 2 levels. For a noisy patch, we calculate the distance between it and all clustering centers of groups. Then we construct the fractional-order patch using the corresponding fractional-order parameter of the nearest patch group. In the sparse coding and dictionary learning stage, we calculate sparse coefficients for fractional-order patches and update dictionary atoms. Finally, the denoised image can be reconstructed by overlapping image patches. Therefore, the proposed algorithm can be summarized in Algorithm 1.

IV. Experimental Results and Analysis

To demonstrate the superior performance of the proposed FSR model, in this section, we presented extensive experiment results on five typical images with simple and complex textures with various noise levels. We compared the proposed FSR model with three current state-of-the-art image denoising approaches: the conventional sparse representation (SR) denoising model [4], nonlocally centralized sparse representation (NCSR) denoising model [16] and the BM3D image denoising with shaped-adaptive principal component analysis (SAPCA-BM3D) algorithm [3] (it is an enhanced version of BM3D denoising in which local spatial adaptation is achieved by shaped-adaptive principal component analysis (PCA)).

Algorithm 1 Fractional-order Image Denoising Algorithm

Input: Noisy image $\boldsymbol{y}$.
Output: Denoised image $\widehat{\boldsymbol{x}}^{(t+1)}$.
Initialization: Set $\widehat{\boldsymbol{x}}^{(0)} = \boldsymbol{y}$, $D^{(0)} =$ DCT dictionary.
while not converge **do**
1. Construct novel fractional-order sample space.
1) Decompose the noisy image into frequency sub-spaces using the discrete Haar wavelet transform.
2) Construct the novel fractional-order sample space using adaptive fractional-order parameters.
2. Sparse coding and dictionary learning.
1) Calculate sparse coefficient for each fractional-order patch.
2) Update atom $d_l,\ l = 1, 2, \ldots, k$ by
a) Find the set of patches: $\omega_l = \{(i,j) | \boldsymbol{\alpha}_{ij}(l) \neq 0\}$ and calculate the error matrix by
$\boldsymbol{E}_l = \{\boldsymbol{e}_{ij}^l | \boldsymbol{e}_{ij}^l = (\widehat{\boldsymbol{x}}_{ij}^{(t+1)})^{\boldsymbol{\eta}_i} - \sum_{m \neq l} d_m \boldsymbol{\alpha}_{ij}(m)\}_{(i,j) \in \omega_l}$
b) Update the dictionary atom and the coefficient values.
3. If $\|\widehat{\boldsymbol{x}}^{(t+1)} - \widehat{\boldsymbol{x}}^{(t)}\|_2^2 \leq \varepsilon$, break.
end while

In the proposed FSR model, 40 000 overlapping image patches of size 8×8 were directly extracted from noisy images and were clustered into 10 groups. The fraction-order patches in each cluster were constructed using their corresponding fraction-order parameters in the discrete Haar wavelet domain. Then we trained a redundant dictionary of size 64×256 based on the initial DCT dictionary. Finally, the denoised patches were averaged using the Lagrange multiplier $\lambda = 30/\sigma$, as described in (4). The SR denoising model trained a dictionary based on the 40 000 image patches of size 8×8 extracted from the noisy images using the same parameters as those of the proposed FSR model. The NCSR denoising model clustered patches into 70 groups and computed the sparse coefficients using iterative shrinkage algorithm (parameters details in [16]). The SAPCA-BM3D algorithm was applied to the noisy image using a shape-adaptive transform basis. In this method, similar blocks were found in a 25×25 local neighborhood using block matching with the maximum d-distance τ = 2500. In the collaborative filtering stage, PCA was employed as part of the 3D transform.

A. Comparison of Denoising Performance

To demonstrate the denoising performance of the proposed FSR model, we performed a set of experiments on five typical images: Lena, Barbara, Boat, House and Pepper. The experiments were performed 10 times, and the average results were reported for five noise levels σ = 5, 10, 20, 35, 50. We evaluated the quality of the denoised images using the popular PSNR (in dB) and structural similarity index measurement (SSIM).

The PSNR results are listed in Table I. As seen in Table I, the proposed FSR model obtains excellent denoising performance, and 72 % PSNR results are the best among four comparison approaches. Although the SAPCA-BM3D algorithm obtains 20 % of the best PSNR results, the approach is very time-consuming as described in Section IV-B. Note that, with little disturb of the noise, most textures are preserved

at the lowest noise level $\sigma = 5$. For the smooth Pepper image, the SAPCA-BM3D algorithm can sufficiently exploit the self-similarities of patches. Hence, the SAPCA-BM3D result (38.30 dB) is superior to sparse representation-based results, i.e., the SR result (37.79 dB), the NCSR result (38.06 dB) and the proposed FSR result (37.78 dB).

TABLE I
AVERAGE DENOISING RESULTS (PSNR IN DB) FOR VARIOUS NOISY LEVELS, OBTAINED BY THE SAPCA-BM3D, SR, NCSR AND FSR MODEL. THE BEST RESULT IS HIGHLIGHTED.

Method	σ	PSNR (dB)				
		Lena (512^2)	Barbara (512^2)	Boat (512^2)	House (256^2)	Pepper (512^2)
SAPCA-BM3D	5	38.86	38.38	37.50	40.13	**38.30**
	10	36.07	35.07	34.10	37.06	34.94
	20	33.20	31.97	31.02	**34.03**	31.61
	35	30.55	**28.97**	28.42	31.36	28.51
	50	**29.07**	**27.51**	26.89	29.53	26.94
SR	5	38.61	38.27	37.18	39.65	37.79
	10	35.56	34.65	33.73	36.26	34.84
	20	32.36	30.92	30.42	33.33	32.29
	35	29.59	27.69	27.61	30.41	29.94
	50	27.83	25.42	25.93	28.03	27.99
NCSR	5	38.70	38.36	37.35	39.91	38.06
	10	35.81	34.98	33.90	36.80	34.66
	20	32.92	31.72	30.74	33.97	31.26
	35	30.56	28.69	28.14	**31.49**	28.30
	50	28.89	27.10	26.60	**29.63**	26.53
FSR	5	**40.55**	**39.86**	**37.57**	**41.19**	37.78
	10	**37.13**	**36.82**	**35.68**	**37.12**	**36.38**
	20	**33.61**	**32.56**	**32.53**	33.69	**33.98**
	35	**30.78**	28.65	**29.67**	31.00	**31.46**
	50	28.93	26.40	**27.87**	29.11	**29.69**

To better visualize the experimental comparison, Fig. 6(a) presents the average PSNR results over all test images for each noise level. Note that the FSR model leads to the best denoising performance on all noise levels. Although the FSR does not exploit image self-similarities as the SAPCA-BM3D algorithm does, the FSR outperforms SAPCA-BM3D by 0.96 dB. The FSR model outperforms other two sparse representation-based models: the SR model by up to 1.35 dB and the NCSR model by up to 1.0 dB, respectively. Furthermore, the NCSR model performs almost the same as the SAPCA-BM3D algorithm. In terms of the PSNR standard deviation, as depicted in Fig. 6(b), the FSR model improves the robustness for four out of five noise levels. Due to the weak capacity of the FSR model for the smooth Pepper image, the FSR model falls behind other three models at the low noise level $\sigma = 5$.

For further evaluation of the denoising performance, we calculated the SSIM results of the denoised images. Table II shows that 56 % of the SSIM results obtained by our proposed FSR model are higher than those obtained by other three approaches. Note that, for the smooth Pepper image, the average SSIM result is 0.87 for the SAPCA-BM3D algorithm, 0.84 for the SR model, 0.88 for the NCSR model and 0.85 for the FSR model. Thus, for a smooth image, the NCSR model achieves a better result than other two sparse representation models (e.g., SR/FSR). However, the proposed FSR model outperforms best in most cases.

TABLE II
AVERAGE DENOISING RESULTS (SSIM) FOR VARIOUS NOISY LEVELS, OBTAINED BY THE SAPCA-BM3D, SR, NCSR AND FSR MODEL. THE BEST RESULT IS HIGHLIGHTED.

Method	σ	SSIMs				
		Lena (512^2)	Barbara (512^2)	Boat (512^2)	House (256^2)	Pepper (512^2)
SAPCA-BM3D	5	0.94	0.96	0.93	0.95	0.95
	10	0.91	0.94	0.88	0.92	0.92
	20	0.87	0.90	0.82	0.87	0.88
	35	0.82	0.84	0.75	0.83	0.83
	50	**0.79**	0.79	0.70	0.81	0.79
SR	5	0.95	0.96	0.93	0.96	0.93
	10	0.91	0.94	0.88	0.91	0.88
	20	0.86	0.88	0.80	0.86	0.84
	35	0.81	0.79	0.72	0.82	0.80
	50	0.76	0.71	0.66	0.77	0.76
NCSR	5	0.95	0.94	**0.94**	0.96	**0.95**
	10	0.92	0.94	0.88	0.92	**0.92**
	20	0.87	0.90	0.81	**0.87**	**0.88**
	35	**0.83**	0.85	0.74	**0.84**	**0.84**
	50	0.72	0.78	0.69	**0.82**	**0.79**
FSR	5	**0.96**	**0.97**	0.93	**0.97**	0.91
	10	**0.93**	**0.95**	**0.91**	**0.92**	0.90
	20	**0.87**	**0.92**	**0.86**	0.86	0.86
	35	0.80	**0.85**	**0.78**	0.83	0.80
	50	0.73	**0.79**	**0.72**	0.78	0.77

Figs. 7(a) and 7(b) show the visual comparison of the average SSIM and their standard deviation. Fig. 7(a) shows that the FSR is slightly better than other three methods for low and moderate noise levels ($\sigma = 5, 10, 20$). Because the SAPCA-BM3D algorithm and NCSR model utilize the structural similarity among numerous similar image patches, the corresponding results are superior to the FSR results at the high noise levels ($\sigma = 35, 50$). In terms of the SSIM standard deviation, as depicted in Fig. 7(b), the proposed FSR model improves the robustness except at the low noise level ($\sigma = 5$).

Figs. 8–10 illustrate the visual comparison of denoised results for three typical images (Pepper, Boat and Barbara) at moderate noise levels ($\sigma = 20$). Figs. 8(a)−8(e) show the denoised images, and Figs. 8(g)−8(j) show the corresponding noisy components. It can be observed from Fig. 8 that the FSR model delivers the best visual quality. The SAPCA-BM3D and SR methods tend to suffer from noticeable artifacts around the smooth areas, as shown in Fig. 9. Furthermore, for regions with complex textures, the superiority of the FSR model is visually justified by Fig. 10.

Thus, the FSR model can both preserve the sharpness of edges and suppress undesirable artifacts. Based on the above discussion in this sub-section, we can conclude that the FSR

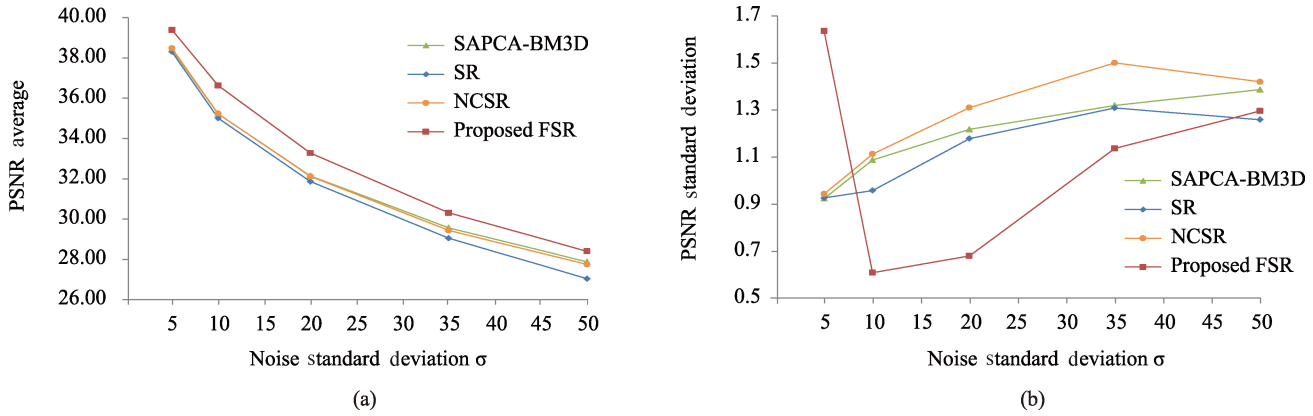


Fig. 6. Comparison of the average and the standard deviations of PSNR on various noise standard deviations.

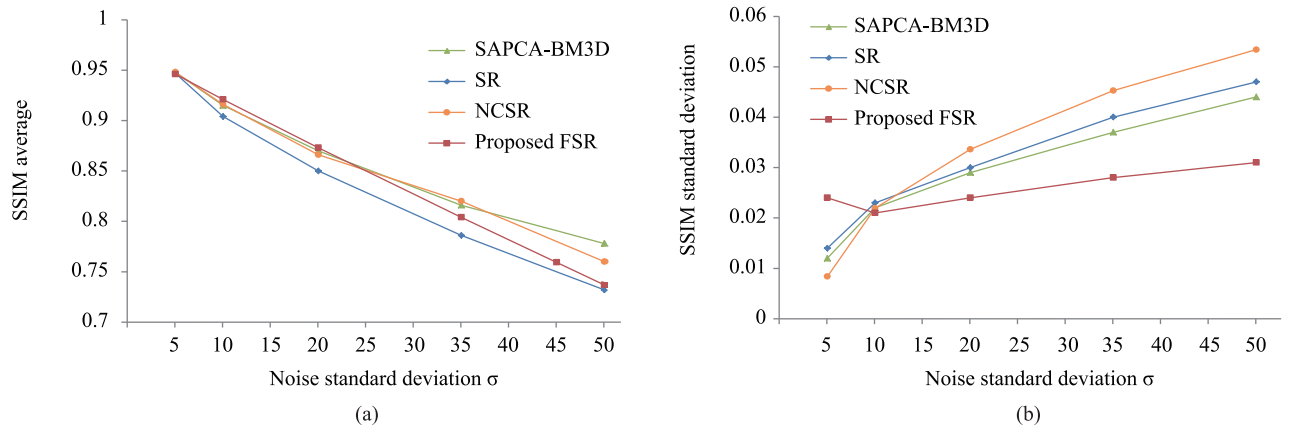


Fig. 7. Comparison of the average and the standard deviations of SSIM for various noise standard deviations.

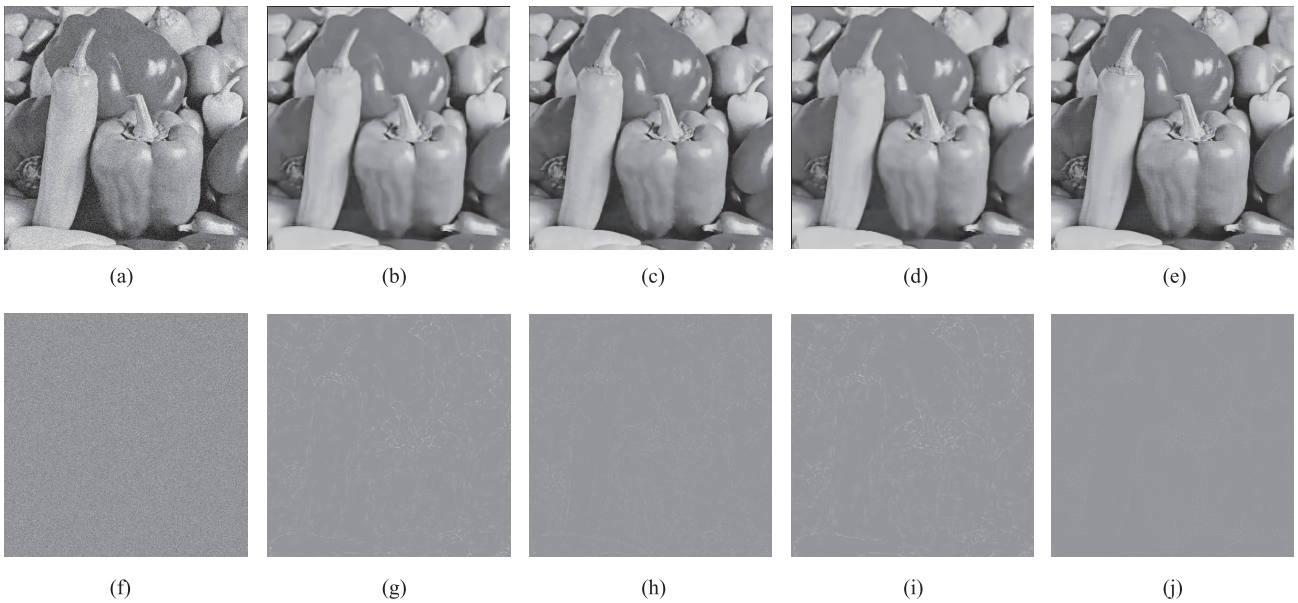


Fig. 8. Denoising performance comparison on the Pepper image with moderate noise corruption. (a) Noisy image ($\sigma = 20$); denoised images by (b) SAPCA-BM3D (PSNR = 31.61 dB, SSIM = 0.88); (c) SR (PSNR = 32.29 dB, SSIM=0.84); (d) NCSR (PSNR = 31.26 dB, SSIM = 0.88); (e) FSR (PSNR = 33.98 dB, SSIM = 0.86); (f) real noisy component; corresponding noisy components by (g) SAPCA-BM3D; (h) SR; (i) NCSR; (j) FSR.

(a) (b) (c) (d) (e)

Fig. 9. Denoising performance comparison on the Boat image with moderate noise corruption ($\sigma = 20$). (a) Original image; denoised images by (b) SAPCA-BM3D (PSNR = 31.02 dB, SSIM = 0.82); (c) SR (PSNR = 30.42 dB, SSIM = 0.80); (d) NCSR (PSNR = 30.74 dB, SSIM = 0.81); (e) FSR (PSNR = 32.53 dB, SSIM = 0.86).

(a) (b) (c) (d) (e)

Fig. 10. Denoising performance comparison on the Barbara image with moderate noise corruption ($\sigma = 20$). (a) Original image; denoised images by (b) SAPCA-BM3D (PSNR = 31.97 dB, SSIM = 0.90); (c) SR (PSNR = 30.92 dB, SSIM = 0.88); (d) NCSR (PSNR = 31.72 dB, SSIM = 0.90); (e) FSR (PSNR = 32.56 dB, SSIM = 0.92).

model improves the denoising performance and robustness in terms of the PSNR and SSIM.

B. Comparison of Computational Efficiency

In this section, we discussed the computational cost of our FSR model in comparison to the SAPCA-BM3D algorithm and the sparse representation-based models (e.g., SR, NCSR). To make a fair comparison, the FSR model and the sparse representation-based models used the same stopping scalar controlling the convergence of the iterative. All the experiments were implemented on a personal computer (Intel i7-4790 CPU, 3.60 GHz, 16 GB RAM) with MATLAB 8.3 on 64-bit Windows 8. The experiments were repeated 10 times, and the results were reported with the average running time values for different noise levels $\sigma = 5$, 10, 20, 35, 50.

The computational cost of the proposed FSR model with comparison to other competing methods is reported in Table III. We can see that the SAPCA-BM3D algorithm and the NCSR model are time-consuming. For the SAPCA-BM3D algorithm, there is little difference in running time among the five noise levels. Note that the running time is directly proportional to the size of the image: 331.55 s for the House image of size 256×256 and 1329.63 s for the Lena image of size 512×512 at $\sigma = 5$. And the NCSR model requires more time for image denoising when the noise level increases. By contrast, the running time for both FSR and SR model declines with the increasing of noise, because the seriously corruption of noisy images at heavy noise levels $\sigma = 35, 50$, these models can only learn few image texture details during the dictionary learning stage and converge quickly. The FSR model is more than 100 times faster than the SAPCA-BM3D and is over 30

TABLE III

TIME CONSUMPTION RESULTS (IN SECONDS) FOR VARIOUS NOISY LEVELS OBTAINED BY THE SAPCA-BM3D, SR, NCSR AND FSR MODEL. THE LAST COLUMN PRESENTS THE AVERAGE RUNNING TIME RESULTS OVER ALL TEST IMAGES. THE BEST RESULT IS HIGHLIGHTED.

Method	σ	Running time (s)					Average
		Lena (512^2)	Barbara (512^2)	Boat (512^2)	House (256^2)	Pepper (512^2)	
SAPCA-BM3D	5	1329.63	1403.72	1375.73	331.55	1355.65	966.88
	10	1331.62	1352.91	1387.30	332.89	1354.90	961.60
	20	1334.86	1348.03	1377.07	333.39	1353.00	961.06
	35	1339.29	1355.45	1384.13	338.13	1357.21	968.20
	50	1318.81	1336.48	1386.97	341.64	1330.25	960.69
SR	5	48.20	99.43	84.30	51.97	52.94	67.37
	10	16.82	29.42	25.68	17.14	16.05	21.02
	20	6.76	9.62	9.54	6.21	6.38	7.70
	35	3.50	4.29	4.60	2.88	3.24	3.70
	50	2.78	2.54	2.79	1.77	2.08	2.39
NCSR	5	677.65	687.39	699.47	154.73	674.60	578.77
	10	650.09	676.66	645.27	150.54	653.84	555.28
	20	859.17	804.57	755.78	189.36	797.76	681.33
	35	1859.78	1643.56	1598.92	424.71	1728.76	1451.15
	50	1891.67	1708.95	1757.39	431.92	1769.44	1511.87
FSR	5	**12.88**	**23.88**	**23.17**	**10.41**	**13.73**	**16.81**
	10	**4.34**	**6.35**	**6.74**	**3.72**	**4.27**	**5.08**
	20	**2.66**	**3.17**	**2.59**	**1.81**	**2.48**	**2.54**
	35	**2.28**	**2.42**	**2.41**	**1.65**	**2.01**	**2.15**
	50	**2.14**	**2.15**	**2.20**	**1.26**	**1.54**	**1.86**

times faster than other sparse representation-based models. Consequently, the running time of the FSR model is significantly reduced, and the FSR model exhibits the best computational efficiency among the four approaches.

V. Conclusion

To obtain more accurate sparse coefficients and optimal dictionary, we presented a new model for image denoising called the fractional-order sparse representation (FSR) model. The proposed FSR model has been evaluated on the basis of five typical images with various noise levels. A series experimental results show that, in most cases, the FSR model significantly outperforms the SAPCA-BM3D and other two sparse representation-based models (e.g., SR, NCSR). Furthermore, the FSR model performs with remarkable computational efficiency. Therefore, we can conclude that the proposed FSR model achieves state-of-the-art performance in terms of both denoising effectiveness and computational efficiency. However, the fractional-order parameters are learned from external images database. It remains unclear how to set fractional-order parameters automatically, which is out of scope of this paper and will lead to our future work.

References

[1] M. Afonso and J. M. Sanches, "Image reconstruction under multiplicative speckle noise using total variation," *Neurocomputing*, vol. 150, pp. 200−213, Feb. 2015.

[2] Z. G. Sun, S. C. Chen, and L. S. Qiao, "A general non-local denoising model using multi-kernel-induced measures," *Pattern Recognit.*, vol. 47, no. 4, pp. 1751−1763, Apr. 2014.

[3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *Proc. Workshop on Signal Processing with Adaptive Sparse Structured Representation*, Saint Malo, France, 2009.

[4] Y. Romano and M. Elad, "Improving K-SVD denoising by post-processing its method-noise," in *Proc. 20th IEEE Int. Conf. Image Processing*, Melbourne, VIC, Australia, 2013, pp. 435−439.

[5] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736−45, Dec. 2006.

[6] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34−81, Feb. 2009.

[7] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1438−1451, Jul. 2009.

[8] M. Zibulevsky and M. Elad, "L1-l2 optimization in signal and image processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 76−88, May 2010.

[9] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proc. IEEE*, vol. 98, no. 6, pp. 948−958, Jun. 2010.

[10] X. Q. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253−276, Jul. 2010.

[11] A. Marquina and S. J. Osher, "Image super-resolution by TV-regularization and Bregman iteration," *J. Sci. Comp.*, vol. 37, no. 3, pp. 367−382, Dec. 2008.

[12] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, New York, NY, USA, 2006, pp. 895−900.

[13] R. Rubinstein, T. Peleg, and M. Elad, "Analysis K-SVD: A dictionary-learning algorithm for the analysis sparse model," *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 661−677, Feb. 2013.

[14] J. C. Yang, Z. W. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution." *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3467−78, 2012.

[15] Y. B. Tang, N. Xu, A. M. Jiang, and C. P. Zhu, "Image denoising via sparse representation using rotational dictionary," *J. Electron. Imag.* vol. 23, no. 5, Article ID 053016, Oct. 2014.

[16] W. S. Dong, L. Zhang, G. M. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1620−1630, Apr. 2013.

[17] W. S. Dong, G. M. Shi, Y. Ma, and X. Li, "Image restoration via simultaneous sparse coding: Where structured sparsity meets Gaussian scale mixture," *Int. J. Comp. Vis.*, vol. 114, no. 2−3, pp. 217−232, Sep. 2015.

[18] N. He, J. B. Wang, L. L. Zhang, and K. Lu, "An improved fractional-order differentiation model for image denoising," *Signal Process.*, vol. 112, pp. 180−188, Jul. 2015.

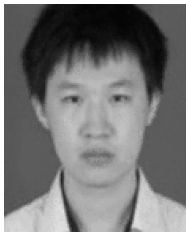
[19] Y. F. Pu, J. L. Zhou, and X. Yuan, "Fractional differential mask: a fractional differential-based approach for multiscale texture enhancement," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 491−511, Feb. 2010.

[20] S. Larnier and R. Mecca, "Fractional-order diffusion for image reconstruction," in *Proc. 2012 IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Kyoto, Japan, 2012, pp. 1057−1060.

[21] Y. H. Yuan, Q. S. Sun, and H. W. Ge, "Fractional-order embedding canonical correlation analysis and its applications to multi-view dimensionality reduction and recognition," *Pattern Recognit.*, vol. 47, no. 3, pp. 1411−1424, Mar. 2014.

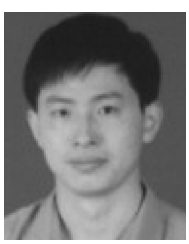
[22] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore: Johns Hopkins University Press, 1996.

Leilei Geng graduated from Nanjing University of Science and Technology (NUST), China, 2016. She received the Ph.D. degree in pattern recognition and intelligence system from NUST, Nanjing, China, in 2016. She is currently a Lecturer at the School of Computer Science and Technology, Shandong University of Finance and Economics. Her research interests include image processing, remote sensing imaging, and pattern recognition.



Zexuan Ji graduated from Nanjing University of Science and Technology (NUST), China, 2012. He received the Ph.D. degree in pattern recognition and intelligence system from NUST, Nanjing, China, 2012. He is currently an Associate Professor at the School of Computer Science and Engineering, NUST. His research interests include medical imaging, image processing and pattern recognition.

Yunhao Yuan received the Ph.D. degree in pattern recognition and intelligence system from Nanjing University of Science and Technology (NUST), China, in 2013. He is currently an Associate Professor at the School of Internet of Things Engineering, Jiangnan University (JN), China. He has published more than 20 scientific papers. He is currently a member of International Society of Information Fusion (ISIF). His research interests include pattern recognition, image processing, computer vision, and information fusion.



Yilong Yin received the Ph.D. degree in 2002 from Jilin University and worked as a post-doctoral fellow in the Department of Electronic Science and Engineering, Nanjing University. He is currently a Professor at the School of Computer Science and Technology, Shandong University of Finance and Economics, China. His research interests include machine learning, data mining, and biometrics.