Nonlocally Centralized Sparse Representation for Image Restoration

Weisheng Dong, Lei Zhang, *Member, IEEE*, Guangming Shi, *Senior Member, IEEE*, and Xin Li, *Senior Member, IEEE*

***Abstract*— Sparse representation models code an image patch as a linear combination of a few atoms chosen out from an over-complete dictionary, and they have shown promising results in various image restoration applications. However, due to the degradation of the observed image (e.g., noisy, blurred, and/or down-sampled), the sparse representations by conventional models may not be accurate enough for a faithful reconstruction of the original image. To improve the performance of sparse representation-based image restoration, in this paper the concept of sparse coding noise is introduced, and the goal of image restoration turns to how to suppress the sparse coding noise. To this end, we exploit the image nonlocal self-similarity to obtain good estimates of the sparse coding coefficients of the original image, and then centralize the sparse coding coefficients of the observed image to those estimates. The so-called nonlocally centralized sparse representation (NCSR) model is as simple as the standard sparse representation model, while our extensive experiments on various types of image restoration problems, including denoising, deblurring and super-resolution, validate the generality and state-of-the-art performance of the proposed NCSR algorithm.**

***Index Terms*— Image restoration, nonlocal similarity, sparse representation.**

I. Introduction

RECONSTRUCTING a high quality image from one or several of its degraded (e.g., noisy, blurred, and/or down-sampled) versions has many important applications, such as medical imaging, remote sensing, surveillance, and entertainment, etc. For an observed image $\boldsymbol{y}$, the problem of image restoration (IR) can be generally formulated by

$$\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \upsilon \tag{1}$$

Manuscript received January 6, 2012; revised October 20, 2012; accepted October 28, 2012. Date of publication December 21, 2012; date of current version February 12, 2013. This work was supported in part by Major State Basic Research Development Program of China (973 Program) under Grant 2013CB329402, the Natural Science Foundation of China under Grant 61033004, Grant 61227004, and Grant 61100154, the Fundamental Research Funds of the Central Universities of China under Grant K50510020003, and the Hong Kong RGC General Research Fund under Grant PolyU 5375/09E. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jose M. Bioucas-Dias.

W. Dong and G. Shi are with the Key Laboratory of Intelligent Perception and Image Understanding of Education, School of Electronic Engineering, Xidian University, Xi'an 710071, China (e-mail: wsdong@mail.xidian.edu.cn; gmshi@xidian.edu.cn).

L. Zhang is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong (e-mail: cslzhang@comp.polyu.edu.hk).

X. Li is with the Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506 (e-mail: Xin.Li@mail.wvu.edu).

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/TIP.2012.2235847

where $\mathbf{H}$ is a degradation matrix, $\boldsymbol{x}$ is the original image vector and υ is the additive noise vector. With different settings of matrix $\mathbf{H}$, Eq. (1) can represent different IR problems; for example, image denoising when $\mathbf{H}$ is an identity matrix, image deblurring when $\mathbf{H}$ is a blurring operator, image super-resolution when $\mathbf{H}$ is a composite operator of blurring and down-sampling, and compressive sensing when $\mathbf{H}$ is a random projection matrix [1]–[3]. In the past decades, extensive studies have been conducted on developing various IR approaches [4]–[23], [28]. Due to the ill-posed nature of IR, the regularization-based techniques have been widely used by regularizing the solution spaces [5]–[9], [12], [22]. In order for an effective regularizer, it is of great importance to find and model the appropriate prior knowledge of natural images, and various image prior models have been developed [5]–[8], [14], [17], [18], [22].

The classic regularization models, such as the quadratic Tikhonov regularization [8] and the TV regularization [5]–[7] are effective in removing the noise artifacts but tend to over-smooth the images due to the piecewise constant assumption. As an alternative, in recent years the sparsity-based regularization [9]–[23] has led to promising results for various image restoration problems [1]–[3], [16]–[23]. Mathematically, the sparse representation model assumes that a signal $\boldsymbol{x} \in \Re^N$ can be represented as $\boldsymbol{x} \approx \boldsymbol{\Phi\alpha}$, where $\Phi \in \Re^{n\times M}(N < M)$ is an over-complete dictionary, and most entries of the coding vector $\boldsymbol{\alpha}$ are zero or close to zero. The sparse decomposition of $\boldsymbol{x}$ can be obtained by solving an l_0-minimization problem, formulated as $\boldsymbol{\alpha}_{\boldsymbol{x}} = \arg\min_{\boldsymbol{\alpha}} ||\boldsymbol{\alpha}||_0, s.t. ||\boldsymbol{x} - \boldsymbol{\Phi\alpha}||_2 \leq \varepsilon$, where $\| \bullet ||_0$ is a pseudo norm that counts the number of non-zero entries in $\boldsymbol{\alpha}$, and ε is a small constant controlling the approximation error. Since l_0-minimization is an NP-hard combinatorial optimization problem, it is often relaxed to the convex l_1-minimization. The l_1-norm based sparse coding problem can be generally formulated in the following Lagrangian form:

$$\boldsymbol{\alpha}_{\boldsymbol{x}} = \arg\min_{\boldsymbol{\alpha}} \{||\boldsymbol{x} - \boldsymbol{\Phi\alpha}||_2^2 + \lambda||\boldsymbol{\alpha}||_1\} \tag{2}$$

where constant λ denotes the regularization parameter. With an appropriate selection of the regularization parameter λ, we can get a good balance between the sparse approximation error of $\boldsymbol{x}$ and the sparsity of $\boldsymbol{\alpha}$, and the term "sparse coding" refer to this sparse approximation process of $\boldsymbol{x}$. Many efficient l_1-minimization techniques have been proposed to solve Eq. (2), such as iterative thresholding algorithms [9]–[11] and Bregman split algorithms [24], [25]. In addition, compared with the

Fig. 1. Examples of the sparse coding coefficients by using the KSVD [26] approach. (a) Some natural images. (b) Corresponding distributions of the sparse coding coefficients (associated with the 3rd atom of the dictionary in KSVD) of the patches extracted at each pixel. Note that the coefficients are not randomly distributed but highly correlated.

analytically designed dictionaries (e.g., wavelet/curvelet dictionary), the dictionaries learned from example image patches can improve much the sparse representation performance since they can better characterize the image structures [26], [27].

In the scenario of IR, what we observed is the degraded image signal $\boldsymbol{y}$ via $\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \upsilon$. To recover $\boldsymbol{x}$ from $\boldsymbol{y}$, first $\boldsymbol{y}$ is sparsely coded with respect to Φ by solving the following minimization problem:

$$\boldsymbol{\alpha}_{\mathbf{y}} = \arg\min_{\boldsymbol{\alpha}} \{||\boldsymbol{y} - \mathbf{H}\Phi\boldsymbol{\alpha}||_2^2 + \lambda||\boldsymbol{\alpha}||_1\} \tag{3}$$

and then $\boldsymbol{x}$ is reconstructed by $\hat{\boldsymbol{x}} = \Phi\boldsymbol{\alpha}_{\mathbf{y}}$. Clearly, it is expected that $\boldsymbol{\alpha}_{\boldsymbol{y}}$ could be close enough to $\boldsymbol{\alpha}_{\boldsymbol{x}}$. Due to the degradation of the observed image (e.g., the image is blurry and noisy), however, it is very challenging to recover the true sparse code $\boldsymbol{\alpha}_{\boldsymbol{x}}$ from $\boldsymbol{y}$. Using only the local sparsity constraint $||\boldsymbol{\alpha}||_1$ in Eq. (3) may not lead to an accurate enough image reconstruction. On the other hand, it is known that image sparse coding coefficients $\boldsymbol{\alpha}$ are not randomly distributed due to the local and nonlocal correlations existing in natural images. In Fig. 1, we visualize the sparse coding coefficients of several example images. One can see that the sparse coding coefficients are correlated, while the strong correlations allow us to develop a much more accurate sparse model by exploiting the local and nonlocal redundancies. Indeed, some recent works, such as [17] and [18], are based on such considerations. For example, in [18] a group sparse coding scheme was proposed to code similar patches simultaneously, and it achieves impressive denoising results.

In this paper we improve the sparse representation performance by proposing a *nonlocally centralized sparse representation* (NCSR) model. To faithfully reconstruct the original image, the sparse code $\boldsymbol{\alpha}_{\boldsymbol{y}}$ [refer to Eq. (3)] should be as close as possible to the sparse codes $\boldsymbol{\alpha}_{\boldsymbol{x}}$ [refer to Eq. (2)] of the original image. In other words, the difference $\upsilon_{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_{\boldsymbol{y}} - \boldsymbol{\alpha}_{\boldsymbol{x}}$ (called as sparse coding noise, SCN in short, in this work) should be reduced and hence the quality of reconstructed image $\hat{\boldsymbol{x}} = \Phi\boldsymbol{\alpha}_{\mathbf{y}}$ can be improved because $\hat{\boldsymbol{x}} - \boldsymbol{x} \approx \Phi\boldsymbol{\alpha}_{\boldsymbol{y}} - \Phi\boldsymbol{\alpha}_{\boldsymbol{x}} = \Phi\upsilon_{\boldsymbol{\alpha}}$. To reduce the SCN, we centralize the sparse codes to some good estimation of $\boldsymbol{\alpha}_{\boldsymbol{x}}$. In practice, a good estimation of $\boldsymbol{\alpha}_{\boldsymbol{x}}$ can be obtained by exploiting the rich amount of nonlocal redundancies in the observed image.

The proposed NCSR model can be solved effectively by conventional iterative shrinkage algorithm [9], which allows us to adaptively adjust the regularization parameters from a Bayesian viewpoint. The extensive experiments conducted on typical IR problems, including image denoising, deblurring and super-resolution, demonstrate that the proposed NCSR based IR method can achieve highly competitive performance to state-of-the-art denoising methods (e.g., BM3D [17], [39]–[41], LSSC [18]), and outperforms state-of-the-art image deblurring and super-resolution methods.

The rest of this paper is organized as follows. Section II presents the modeling of NCSR. Section III provides the iterative shrinkage algorithm for solving the NCSR model. Section IV presents extensive experimental results and Section V concludes this paper.

II. Nonlocally Centralized Sparse Representation (NCSR)

Following the notation used in [19], for an image $\boldsymbol{x} \in \Re^N$, let $\boldsymbol{x}_i = \mathbf{R}_i\boldsymbol{x}$ denote an image patch of size $\sqrt{n} \times \sqrt{n}$ extracted at location i, where $\mathbf{R}_i$ is the matrix extracting patch $\boldsymbol{x}_i$ from $\boldsymbol{x}$ at location i. Given an dictionary $\Phi \in \Re^{n\times M}, n \le M$, each patch can be sparsely represented as $\boldsymbol{x}_i \approx \Phi\boldsymbol{\alpha}_{x,i}$ by solving an l_1-minimization problem $\boldsymbol{\alpha}_{\boldsymbol{x},i} = \arg\min_{\boldsymbol{\alpha}_i}\{||\boldsymbol{x}_i - \Phi\boldsymbol{\alpha}_i||_2^2 + \lambda||\boldsymbol{\alpha}_i||_1\}$. Then the entire image $\boldsymbol{x}$ can be represented by the set of sparse codes $\{\boldsymbol{\alpha}_{\boldsymbol{x},i}\}$. The patches can be overlapped to suppress the boundary artifacts, and we obtain a redundant patch-based representation. Reconstructing $\boldsymbol{x}$ from $\{\boldsymbol{\alpha}_{x,i}\}$ is an over-determined system, and a straightforward least-square solution is [19]: $\boldsymbol{x} \approx (\sum_{i=1}^N \mathbf{R}_i^T\mathbf{R}_i)^{-1}\sum_{i=1}^N(\mathbf{R}_i^T\Phi\boldsymbol{\alpha}_{\boldsymbol{x},i})$. For the convenience of expression, we let

$$\boldsymbol{x} \approx \Phi \circ \boldsymbol{\alpha}_{\boldsymbol{x}} = \left(\sum\nolimits_{i=1}^N \mathbf{R}_i^T\mathbf{R}_i\right)^{-1}\sum\nolimits_{i=1}^N\left(\mathbf{R}_i^T\Phi\boldsymbol{\alpha}_{x,i}\right) \tag{4}$$

where $\boldsymbol{\alpha}_{\boldsymbol{x}}$ denotes the concatenation of all $\boldsymbol{\alpha}_{\boldsymbol{x},i}$. The above equation is nothing but telling that the overall image is reconstructed by averaging each reconstructed patch of $\boldsymbol{x}_i$.

In the scenario of image restoration (IR), the observed image is modeled as $\boldsymbol{y} = \mathbf{H}\boldsymbol{x} + \upsilon$. The sparsity-based IR method recovers $\boldsymbol{x}$ from $\boldsymbol{y}$ by solving the following minimization problem:

$$\boldsymbol{\alpha}_y = \arg\min_{\alpha}\{||\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}||_2^2 + \lambda||\boldsymbol{\alpha}||_1\}. \tag{5}$$

The image $\boldsymbol{x}$ is then reconstructed as $\hat{\boldsymbol{x}} = \Phi \circ \boldsymbol{\alpha}_y$.

A. Sparse Coding Noise

In order for an effective IR, the sparse codes $\boldsymbol{\alpha}_{\mathbf{y}}$ obtained by solving the objective function in Eq. (5) are expected to be as close as possible to the true sparse codes $\boldsymbol{\alpha}_{\boldsymbol{x}}$ of the original image $\boldsymbol{x}$. However, due to the degradation of the observed image $\boldsymbol{y}$ (e.g., noisy and blurred), the sparse code $\boldsymbol{\alpha}_{\boldsymbol{y}}$ will deviate from $\boldsymbol{\alpha}_{\boldsymbol{x}}$, and the IR quality depends on the level of the sparse coding noise (SCN), which is defined as the difference between $\boldsymbol{\alpha}_{\boldsymbol{y}}$ and $\boldsymbol{\alpha}_{\boldsymbol{x}}$

$$\upsilon_{\boldsymbol{\alpha}} = \boldsymbol{\alpha}_{\boldsymbol{y}} - \boldsymbol{\alpha}_{\boldsymbol{x}}. \tag{6}$$

To investigate the statistical property of SCN $\upsilon_{\boldsymbol{\alpha}}$, we perform some experiments on typical IR problems. We use the

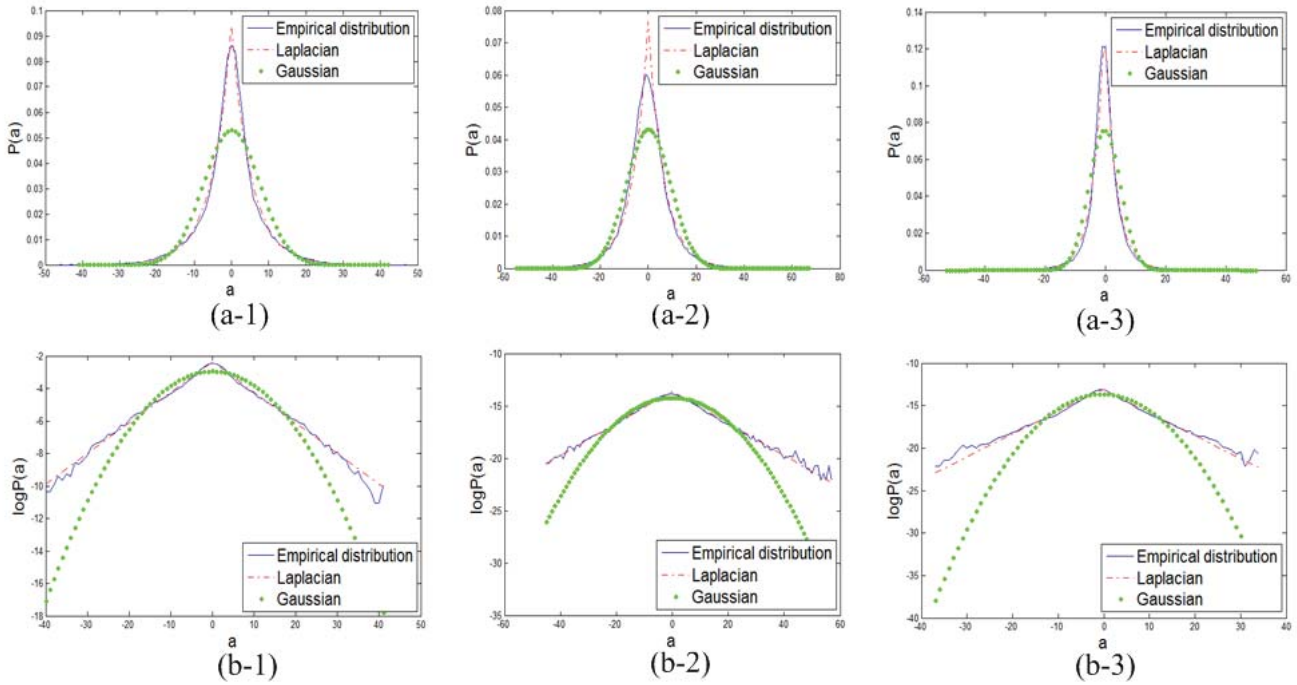

Fig. 2. Distributions of SCN when the *Lena* image is (a-1) noisy, (a-2) noisy and blurred, and (a-3) down-sampled. (b-1)–(b-3) show the same distributions of (a-1)–(a-3) in log domain, respectively.

image *Lena* as an example. In the first experiment, we add Gaussian white noise to the original image $\boldsymbol{x}$ to get the noisy image $\boldsymbol{y}$ (the noise level $\sigma_n = 15$). Then we compute $\boldsymbol{\alpha_x}$ and $\boldsymbol{\alpha_y}$ by solving Eq. (2) and Eq. (5), respectively. The DCT bases are adopted in the experiment. Then the SCN $\upsilon_{\boldsymbol{\alpha}}$ is computed. In Fig. 2(a-1), we plot the distribution of $\upsilon_{\boldsymbol{\alpha}}$ corresponding to the 4$^{\text{th}}$ atom in the dictionary. Similarly, in Fig. 2(a-2) and Fig. 2(a-3) we plot the distributions of $\upsilon_{\boldsymbol{\alpha}}$ when the observed data $\boldsymbol{y}$ is blurred (by a Gaussian blur kernel with standard deviation 1.6) and down-sampled by factor 3 in both horizontal and vertical directions (after blurred by a Gaussian blur kernel with standard deviation 1.6), respectively. We can see that the empirical distributions of SCN $\upsilon_{\boldsymbol{\alpha}}$ can be well characterized by Laplacian distributions, while the Gaussian distributions have much larger fitting errors. To better observe the fitting of the tails, we also plot the these distributions in log domain in Fig. 2(b-1)∼(b-3). This observation motivates us to model $\upsilon_{\boldsymbol{\alpha}}$ with a Laplacian prior, as will be further discussed in Section III-A.

B. Modeling of NCSR

The definition of SCN $\upsilon_{\boldsymbol{\alpha}}$ indicates that by suppressing the SCN $\upsilon_{\boldsymbol{\alpha}}$ we could improve the IR output $\hat{\boldsymbol{x}}$. However, the difficulty lies in that the sparse coding vector $\boldsymbol{\alpha_x}$ is unknown so that $\upsilon_{\boldsymbol{\alpha}}$ cannot be directly measured. Nonetheless, if we could have some reasonably good estimation of $\boldsymbol{\alpha_x}$, denoted by $\boldsymbol{\beta}$, available, then $\boldsymbol{\alpha_y} - \boldsymbol{\beta}$ can be a good estimation of the SCN $\upsilon_{\boldsymbol{\alpha}}$. To suppress $\upsilon_{\boldsymbol{\alpha}}$ and improve the accuracy of $\boldsymbol{\alpha_y}$ and thus further improve the objective function of Eq. (5), we can propose the following centralized sparse representation (CSR) model [22]:

$$\boldsymbol{\alpha_y} = \arg\min_{\alpha} \left\{ \|\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}\|_2^2 + \lambda \sum_i ||\boldsymbol{\alpha}_i||_1 + \gamma \sum_i ||\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i||_p \right\} \quad (7)$$

where $\boldsymbol{\beta}_i$ is some good estimation of $\boldsymbol{\alpha}_i$, γ is the regularization parameter and p can be 1 or 2. In the above CSR model, while enforcing the sparsity of coding coefficients $\boldsymbol{\alpha}_i$, the sparse codes are also centralized to some estimate of $\boldsymbol{\alpha_x}$ (i.e., $\boldsymbol{\beta}$) so that SCN $\upsilon_{\boldsymbol{\alpha}}$ can be suppressed.

One important issue of sparsity-based IR is the selection of dictionary Φ. Conventional analytically designed dictionaries, such as DCT, wavelet and curvelet dictionaries, are insufficient to characterize the so many complex structures of natural images. The universal dictionaries learned from example image patches by using algorithms such as KSVD [26] can better adapt to local image structures. In general the learned dictionaries are required to be very redundant such that they can represent various image local structures. However, it has been shown that sparse coding with an overcomplete dictionary is unstable [42], especially in the scenario of image restoration. In our previous work [21], we cluster the training patches extracted from a set of example images into K clusters, and learn a PCA sub-dictionary for each cluster. Then for a given patch, one compact PCA sub-dictionary is adaptively selected to code it, leading to a more stable and sparser representation, and consequently better image restoration results. In this paper, we adopt this adaptive sparse domain selection strategy but learn the sub-dictionaries from the given image itself instead of the example images.

We extract image patches from image $\boldsymbol{x}$ and cluster the patches into K clusters (typically $K = 70$) by using the K-means clustering method. Since the patches in a cluster are similar to each other, there is no need to learn an over-complete dictionary for each cluster. Therefore, for each cluster we learn a dictionary of PCA bases and use this compact PCA dictionary to code the patches in this cluster. (For the details of PCA dictionary learning, please refer to [21].) These K PCA sub-dictionaries construct a large over-complete dictionary to characterize all the possible local structures of natural images.

In the conventional sparse representation models as well as the model in Eq. (7), the local sparsity term $||\boldsymbol{\alpha}_i||_1$ is used to ensure that only a small number of atoms are selected from the over-complete dictionary Φ to represent the input image patch. In our algorithm (please refer to **Algorithm 1** in Section III-C), for each patch to be coded, we adaptively select one sub-dictionary from the trained K PCA sub-dictionaries to code it. This actually enforces the coding coefficients of this patch over the other sub-dictionaries to be 0, leading to a very sparse representation of the given patch. In other words, our algorithm will naturally ensure the sparsity of the coding coefficients, and thus the local sparsity regularization term $||\boldsymbol{\alpha}_i||_1$ can be removed. Hence we propose the following sparse coding model:

$$\boldsymbol{\alpha}_{\boldsymbol{y}} = \arg\min_{\alpha} \left\{ \|\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}\|_2^2 + \lambda \sum_i ||\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i||_p \right\}. \quad (8)$$

There is only one regularization term $||\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i||_p$ in the above model. In the case that $p = 1$, and the estimate $\boldsymbol{\beta}_i$ is obtained by using the nonlocal redundancy of natural images, this regularization term will become a nonlocally centralized sparsity term, and we call this model *nonlocally centralized sparse representation* (NCSR). Next let's discuss how to obtain a good estimation $\boldsymbol{\beta}_i$ of the unknown sparse coding vectors $\boldsymbol{\alpha}_{\boldsymbol{i}}$.

C. Nonlocal Estimate of Unknown Sparse Code

Generally, there can be various ways to make an estimate of $\boldsymbol{\alpha}_{\boldsymbol{x}}$, depending on how much the prior knowledge of $\boldsymbol{\alpha}_{\boldsymbol{x}}$ we have. If we have many training images that are similar to the original image $\boldsymbol{x}$, we could learn the estimate $\boldsymbol{\beta}$ of $\boldsymbol{\alpha}_{\boldsymbol{x}}$ from the training set. However, in many practical situations the training images are simply not available. On the other hand, the strong nonlocal correlation between the sparse coding coefficients, as shown in Fig. 1, allows us to learn the estimate $\boldsymbol{\beta}$ from the input data. Based on the fact that natural images often contain repetitive structures, i.e., the rich amount of nonlocal redundancies [30], we search the nonlocal similar patches to the given patch i in a large window centered at pixel i. For higher performance, the search of similar patches can also be carried out across different scales at the expense of higher computational complexity, as shown in [31]. Then a good estimation of $\boldsymbol{\alpha}_i$, i.e., $\boldsymbol{\beta}_i$, can be computed as the weighted average of those sparse codes associated with the nonlocal similar patches (including patch i) to patch i. For each patch $\boldsymbol{x}_i$, we have a set of its similar patches, denoted by $\boldsymbol{\Omega}_i$. Finally $\boldsymbol{\beta}_i$ can be computed from the sparse codes of the patches within $\boldsymbol{\Omega}_i$.

Denote by $\boldsymbol{\alpha}_{i,q}$ the sparse codes of patch $\boldsymbol{x}_{i,q}$ within set $\boldsymbol{\Omega}_i$. Then $\boldsymbol{\beta}_i$ can be computed as the weighted average of $\boldsymbol{\alpha}_{i,q}$

$$\boldsymbol{\beta}_i = \sum\nolimits_{q \in \boldsymbol{\Omega}_i} \omega_{i,q} \boldsymbol{\alpha}_{i,q} \quad (9)$$

where $\omega_{i,q}$ is the weight. Similar to the nonlocal means approach [30], we set the weights to be inversely proportional to the distance between patches $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i,q}$

$$\omega_{i,q} = \frac{1}{W} \exp(-||\hat{\boldsymbol{x}}_i - \hat{\boldsymbol{x}}_{i,q}||_2^2 / h) \quad (10)$$

where $\hat{\boldsymbol{x}}_i = \Phi\hat{\boldsymbol{\alpha}}_i$ and $\hat{\boldsymbol{x}}_{i,q} = \Phi\hat{\boldsymbol{\alpha}}_{i,q}$ are the estimates of the patches $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i,q}$, h is a pre-determined scalar and W is the normalization factor. In the case of orthogonal dictionaries (e.g., the local PCA dictionaries used in this work), the sparse codes $\hat{\boldsymbol{\alpha}}_i$ and $\hat{\boldsymbol{\alpha}}_{i,q}$ can be easily computed as $\hat{\boldsymbol{\alpha}}_i = \Phi^T\hat{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{\alpha}}_{i,q} = \Phi^T\hat{\boldsymbol{x}}_{i,q}$. Our experimental results show that by exploiting the nonlocal redundancies of natural images, we are able to achieve good estimation of the unknown sparse vectors $\boldsymbol{\alpha}_i$, and the NCSR model of Eq. (8) can significantly improve the performance of the sparsity-based IR results.

Eq. (8) can be solved iteratively. We first initialize $\boldsymbol{\beta}_i$ as $\mathbf{0}$, i.e., $\boldsymbol{\beta}_i^{(-1)} = \mathbf{0}$, and solve for the sparse coding vector, denoted by $\boldsymbol{\alpha}_{\boldsymbol{y}}^{(0)}$, using some standard sparse coding algorithm. Then we can get the initial estimation of $\boldsymbol{x}$, denoted by $\boldsymbol{x}^{(0)}$, via $\boldsymbol{x}^{(0)} = \Phi \circ \boldsymbol{\alpha}_{\boldsymbol{y}}^{(0)}$. Based on $\boldsymbol{x}^{(0)}$, we search for the similar patches to each patch i, and hence the nonlocal estimate of $\boldsymbol{\beta}_i$ can be updated using Eqs. (9) and (10). The updated estimation of $\boldsymbol{\alpha}_{\boldsymbol{x}}$, denoted by $\boldsymbol{\beta}_i^{(0)}$, will then be used to improve the accuracy of the sparse codes and thus improve the IR quality. Such a procedure is iterated until convergence. In the l^{th} iteration, the sparse vector is obtained by solving the following minimization problem

$$\boldsymbol{\alpha}_y^{(l)} = \arg\min_{\alpha} \left\{ ||\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}||_2^2 + \lambda \sum_i ||\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i^{(l)}||_p \right\}. \quad (11)$$

The restored image is then updated as $\hat{\boldsymbol{x}}^{(l)} = \Phi \circ \boldsymbol{\alpha}_{\boldsymbol{y}}^{(l)}$. In the above iterative process, the accuracy of sparse coding coefficient $\boldsymbol{\alpha}_{\boldsymbol{y}}^{(l)}$ is gradually improved, which in turn improves the accuracy of $\boldsymbol{\beta}_i$. The improved $\boldsymbol{\beta}_i$ are then used to improve the accuracy of $\boldsymbol{\alpha}_{\boldsymbol{y}}$, and so on. Finally, the desired sparse code vector $\boldsymbol{\alpha}_{\boldsymbol{y}}$ is obtained when the alternative optimization process falls into a local minimum. The detailed algorithm will be presented in Section III.

III. Algorithm of NCSR

A. Parameters Determination

In Eq. (8) or Eq. (11) the parameter λ that balances the fidelity term and the centralized sparsity term should be adaptively determined for better IR performance. In this subsection we provide a Bayesian interpretation of the NCSR model, which also provides us an explicit way to set the regularization parameter λ. In the literature of wavelet denoising, the connection between *Maximum a Posterior* (MAP) estimator and sparse representation has been established [28], and here

we extend the connection from the local sparsity to nonlocally centralized sparsity.

For the convenience of expression, let's define $\theta = \boldsymbol{\alpha} - \boldsymbol{\beta}$. For a given $\boldsymbol{\beta}$, the MAP estimation of θ can be formulated as

$$\begin{aligned}\boldsymbol{\theta}_y &= \arg\max_{\boldsymbol{\theta}} \log \mathrm{P}(\boldsymbol{\theta}|\boldsymbol{y}) \\ &= \arg\max_{\boldsymbol{\theta}} \{\log \mathrm{P}(\boldsymbol{y}|\boldsymbol{\theta}) + \log \mathrm{P}(\boldsymbol{\theta})\}. \end{aligned} \tag{12}$$

The likelihood term is characterized by the Gaussian distribution

$$\mathrm{P}(\boldsymbol{y}|\boldsymbol{\theta}) = \mathrm{P}(\boldsymbol{y}|\boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{1}{2\sigma_n^2}||\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}||_2^2\right) \tag{13}$$

where θ and $\boldsymbol{\beta}$ are assumed to be independent. In the prior probability $\mathrm{P}(\theta)$, θ reflects the variation of $\boldsymbol{\alpha}$ from its estimation $\boldsymbol{\beta}$. If we take $\boldsymbol{\beta}$ as a very good estimation of the sparse coding coefficient of unknown true signal, then $\theta_y = \alpha_{\boldsymbol{x}} - \boldsymbol{\beta}$ is basically the SCN associated with $\boldsymbol{\alpha}_{\boldsymbol{y}}$, and we have seen in Fig. 2 that the SCN signal can be well characterized by the Laplacian distribution. Thus, we can assume that θ follows i.i.d. Laplacian distribution, and the joint prior distribution $P(\theta)$ can be modeled as

$$\mathrm{P}(\boldsymbol{\theta}) = \prod_i \prod_j \left\{ \frac{1}{\sqrt{2}\sigma_{i,j}} \exp(-\frac{|\boldsymbol{\theta}_i(j)|}{\sigma_{i,j}}) \right\} \tag{14}$$

where $\theta_i(j)$ are the j^{th} elements of θ_i, and $\sigma_{i,j}$ is the standard deviations of $\theta_i(j)$.

Substituting Eqs. (13) and (14) into Eq. (12), we obtain

$$\begin{aligned}\theta_{\boldsymbol{y}} = \arg\min_{\boldsymbol{\theta}} \Bigg\{ &||\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}||_2^2 + 2\sqrt{2}\sigma_n^2 \\ &\times \sum_i \sum_j \frac{1}{\sigma_{i,j}} |\theta_i(j)| \Bigg\}. \end{aligned} \tag{15}$$

Hence, for a given $\boldsymbol{\beta}$ the sparse codes $\boldsymbol{\alpha}$ can then be obtained by minimizing the following objective function

$$\begin{aligned}\boldsymbol{\alpha}_{\boldsymbol{y}} = \arg\min_{\alpha} \Bigg\{ &\|\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}\|_2^2 + 2\sqrt{2}\sigma_n^2 \\ &\times \sum_i \sum_j \frac{1}{\sigma_{i,j}} |\boldsymbol{\alpha}_i(j) - \boldsymbol{\beta}_i(j)| \Bigg\}. \end{aligned} \tag{16}$$

Compared with Eq. (8), we can see that the l_1-norm (i.e., $p = 1$) should be chosen to characterize the SCN term $\boldsymbol{\alpha}_i - \boldsymbol{\beta}_i$. Comparing Eq. (16) with Eq. (8), we have

$$\lambda_{i,j} = \frac{2\sqrt{2}\sigma_n^2}{\sigma_{i,j}}. \tag{17}$$

In order to have robust estimations of $\sigma_{i,j}$, the image nonlocal redundancies can be exploited. In practice, we estimate $\sigma_{i,j}$ using the set of θ_i computed from the nonlocal similar patches. $\lambda_{i,j}$ is then updated with the updated θ in each iteration or in several iterations to save computational cost. Next we present the detailed algorithm of the proposed NCSR scheme.

Algorithm 1 NCSR-Based Image Restoration

1. Initialization:
   (a) Set the initial estimate as $\hat{x} = \boldsymbol{y}$ for image denoising and deblurring, or initialize $\hat{\boldsymbol{x}}$ by bicubic interpolator for image super-resolution;
   (b) Set initial regularization parameter λ and δ;
2. Outer loop (dictionary learning and clustering): iterate on $l = 1, 2, \ldots, L$
   (a) Update the dictionaries $\{\Phi_k\}$ via k-means and PCA;
   (b) Inner loop (clustering): iterate on $j = 1, 2, \ldots, J$
   (I) $\hat{\boldsymbol{x}}^{(j+1/2)} = \hat{x}^{(j)} + \delta\mathbf{H}^T(\boldsymbol{y} - \mathbf{H}\hat{x}^{(j)})$, where δ is the pre-determined constant;
   (II) Compute $\boldsymbol{v}^{(j)} = [\Phi_{k_1}^T \mathbf{R}_1 \hat{x}^{(j+1/2)}, \ldots, \Phi_{k_N}^T \mathbf{R}_N \hat{\boldsymbol{x}}^{(j+1/2)}]$, where Φ_{k_i} is the dictionary assigned to patch $\hat{x}_i = \mathbf{R}_i \hat{x}^{(j+1/2)}$;
   (III) Compute $\boldsymbol{\alpha}_i^{(j+1)}$ using the shrinkage operator given in Eq. (19);
   (IV) If $mod(j, J_0) = 0$ update the parameters $\lambda_{i,j}$ and $\{\boldsymbol{\beta}_i\}$ using Eqs. (17) and (9), respectively;
   (V) Image estimate update: $\hat{\boldsymbol{x}}^{(j+1)} = \Phi \circ \boldsymbol{\alpha}_{\boldsymbol{y}}^{(j+1)}$ using Eq. (4).

B. Iterative Shrinkage Algorithm

As discussed in Section II, we use an iterative algorithm to solve the NCSR objective function in Eqs. (8) or (16). In each iteration, for fixed $\boldsymbol{\beta}_i$ we solve the following l_1-norm minimization problem:

$$\boldsymbol{\alpha}_{\boldsymbol{y}} = \arg\min_{\alpha} \left\{ \|\boldsymbol{y} - \mathbf{H}\Phi \circ \boldsymbol{\alpha}\|_2^2 + \sum_i \sum_j \lambda_{i,j} |\boldsymbol{\alpha}_i(j) - \boldsymbol{\beta}_i(j)| \right\} \tag{18}$$

which is convex and can be solved efficiently. In this paper we adopt the surrogate algorithm in [9] to solve Eq. (18). In the $(l+1)$-*th* iteration, the proposed shrinkage operator for the j^{th} element of $\boldsymbol{\alpha}_i$ is

$$\boldsymbol{\alpha}_i^{(l+1)}(j) = S_\tau(\boldsymbol{v}_{i,j}^{(l)} - \boldsymbol{\beta}_i(j)) + \boldsymbol{\beta}_i(j) \tag{19}$$

where $S_\tau(\cdot)$ is the classic soft-thresholding operator and $\boldsymbol{v}^{(l)} = K^T(\boldsymbol{y} - K \circ \boldsymbol{\alpha}^{(l)})/c + \boldsymbol{\alpha}^{(l)}$, where $K = \mathbf{H}\Phi$, $K^T = \Phi^T \circ \mathbf{H}^T$, $\tau = \lambda_{i,j}/c$, and c is an auxiliary parameter guaranteeing the convexity of the surrogate function. The derivation of the above shrinkage operator follows the standard surrogate algorithm in [9]. The interesting readers may refer to [9] for details.

C. Summary of the Algorithm

As we mentioned in Section II-B, in our NCSR algorithm the adaptive sparse domain selection strategy [21] is used to code each patch. We cluster the patches of image $\boldsymbol{x}$ into K clusters and learn a PCA sub-dictionary Φ_k for each cluster. For a given patch, we first check which cluster it falls into by calculating its distances to means of the clusters, and then select the PCA sub-dictionary of this cluster to code it.

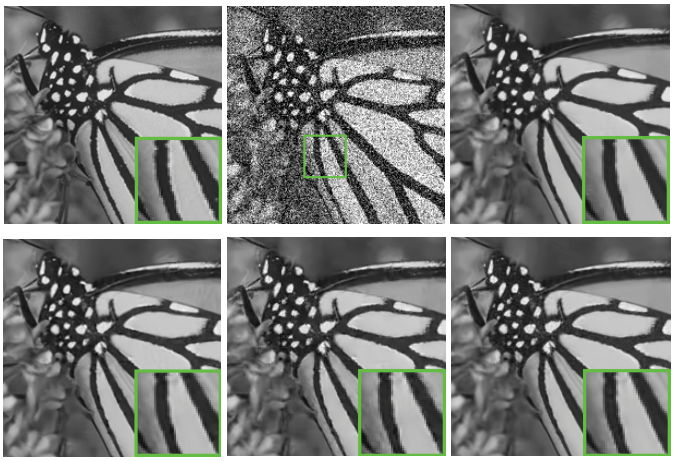

Fig. 3. Denoising performance comparison on the *Monarch* image with moderate noise corruption. From left to right and top to bottom: original image, noisy image ($\sigma = 20$), and denoised images by SAPCA-BM3D [39] (PSNR = **30.91** dB; FSIM = **0.9404**), LSSC [18] (PSNR = 30.58 dB; FSIM = 0.9310), EPLL [33] (PSNR = 30.48 dB; FSIM = 0.9330), and NCSR (PSNR = 30.69 dB; FSIM = 0.9316).

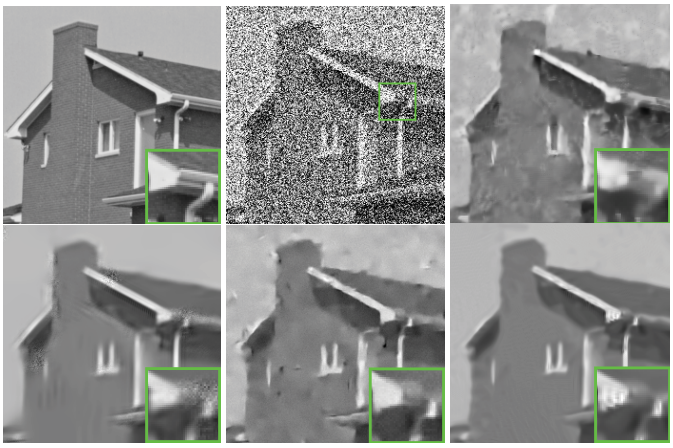

Fig. 4. Denoising performance comparison on the *House* image with strong noise corruption. From left to right and top to bottom: original image, noisy image ($\sigma = 100$), and denoised images by SAPCA-BM3D [39] (PSNR = 25.20 dB; FSIM = 0.8065), LSSC [18] (PSNR = 25.63 dB; FSIM = 0.8017), EPLL [33] (PSNR = 25.44 dB; FSIM = 0.8100), and NCSR (PSNR = **25.65** dB; FSIM = **0.8068**).

The proposed NCSR based IR algorithm is summarized in **Algorithm 1**.

In **Algorithm 1**, for fixed parameters $\lambda_{i,j}$ and $\{\boldsymbol{\beta}_i\}$ the objective function in Eq. (18) is convex and can be efficiently solved by the iterative shrinkage algorithm in the inner loop, and its convergence has been well established in [9]. Since we update the regularization parameter $\lambda_{i,j}$ and $\{\boldsymbol{\beta}_i\}$ in every J_0 iterations after solving a sub-optimization problem, **Algorithm 1** is empirically convergent in general, as those presented in [38].

IV. Experimental Results

To verify the IR performance of the proposed NCSR algorithm we conduct extensive experiments on image denoising, deblurring and super-resolution. The basic parameter setting of NCSR is as follows: the patch size is 7×7 and $K = 70$. For image denoising, $\delta = 0.02$, $L = 3$, and $J = 3$; for image deblurring and super-resolution, $\delta = 2.4$, $L = 5$, and $J = 160$. To evaluate the quality of the restored images, the PSNR and the recently proposed powerful perceptual quality metric FSIM [32] are calculated. Due to the limited page space, we only show part of the results in this paper, and all the experimental results can be downloaded on the website: http://www.comp.polyu.edu.hk/∼cslzhang/NCSR.htm.

A. Image Denoising

We compare the proposed NCSR method with three recently developed state-of-the-art denoising methods, including the shape-adaptive PCA based BM3D (SAPCA-BM3D) [39], [40] (which outperforms the benchmark BM3D algorithm [17]), the learned simultaneously sparse coding (LSSC) method [18] and the expected patch log likelihood (EPLL) based denoising method [33]. A set of 12 natural images commonly used in the literature of image denoising are used for the comparison study. The PSNR results of the test methods are reported in Table I. From Table I, we can see that the proposed NCSR achieves highly competitive denoising performance. In term of average PSNR results, NCSR performs almost the same as LSSC, and is slightly lower than SAPCA-BM3D, which is the best among the competitors.

In Fig. 3 and Fig. 4 we show the denoising results on two typical images with moderate noise corruption and strong noise corruption, respectively. It can be seen that NCSR is very effective in reconstructing both the smooth and the texture/edge regions. When the noise level is not very high, as shown in Fig. 3 ($\sigma = 20$), all the four competing methods can achieve very good denoising outputs. When the noise level is high, as shown in Fig. 4 ($\sigma = 100$), however, the SAPCA-BM3D and EPLL methods tend to generate many visual artifacts. LSSC and NCSR work much better in this case. In particular, the denoised image by the proposed NCSR has much less artifacts than other methods, and is visually more pleasant.

B. Image Deblurring

We applied deblurring methods to both the simulated blurred images and real motion blurred images. In the simulated image deblurring two commonly used blur kernels, i.e., 9×9 uniform blur and 2D Gaussian function (non-truncated) with standard deviation 1.6, are used for simulations. Additive Gaussian noise with noise levels $\sigma_n = \sqrt{2}$ is added to the blurred images. In addition, 6 typical non-blind deblurring image experiments presented in [36] and [41] are conducted for further test. For the real motion blurred images, we borrowed the motion blur kernel estimation method from [34] to estimate the blur kernel and then fed the estimated blur kernel into the NCSR deblurring method. For color images, we only apply the deblurring operation to the luminance component.

We compared the NCSR deblurring method with four state-of-the-art deblurring methods, including the constrained TV deblurring (denoted by FISTA) method [35], the l_0-sparsity based deblurring (denoted by l_0-SPAR) method [36], the

TABLE I

PSNR (dB) RESULTS BY DIFFERENT DENOISING METHODS. IN EACH CELL, THE RESULTS OF THE FOUR DENOISING METHODS ARE REPORTED IN THE FOLLOWING ORDER: TOP LEFT–SAPCA-BM3D [39]. TOP RIGHT–LSSC [18]. BOTTOM LEFT–EPLL [33]. BOTTOM RIGHT–NCSR.

σ	5		10		15		20		50		100	
Lena	**38.86**	38.68	**36.07**	35.83	**34.43**	34.14	**33.20**	32.88	**29.07**	28.95	25.37	**25.96**
	38.52	38.70	35.56	35.81	33.85	34.09	32.60	32.92	28.42	28.89	25.30	25.66
Monarch	**38.69**	38.53	**34.74**	34.48	**32.46**	32.15	**30.92**	30.58	**26.28**	25.59	**22.31**	21.82
	38.22	38.49	34.27	34.57	32.04	32.34	30.48	30.69	25.67	25.68	22.04	22.05
Barbara	38.38	**38.44**	**35.07**	34.95	**33.27**	32.96	**31.97**	31.53	**27.51**	27.13	23.05	**23.56**
	37.56	38.36	33.59	34.98	31.33	33.02	29.75	31.72	24.83	27.10	22.10	23.30
Boat	**37.50**	37.34	**34.10**	33.99	**32.29**	32.17	**31.02**	30.87	**26.89**	26.76	23.71	**23.94**
	36.78	37.35	33.63	33.90	31.89	32.03	30.63	30.74	26.64	26.60	23.78	23.64
C. Man	**38.54**	38.24	**34.52**	34.14	**32.31**	31.96	**30.86**	30.54	**26.59**	26.36	22.91	**23.14**
	38.04	38.17	33.94	34.12	31.73	31.99	30.28	30.48	26.08	26.16	22.87	22.89
Couple	**37.60**	37.41	**34.13**	33.96	**32.20**	32.06	**30.83**	30.70	**26.48**	26.31	23.19	**23.34**
	37.32	37.44	33.78	33.94	31.83	31.95	30.47	30.56	26.22	26.21	**23.34**	23.22
F. Print	36.67	36.71	32.65	32.57	**30.46**	30.31	28.97	28.78	**24.53**	24.21	21.07	21.18
	36.41	**36.81**	32.13	**32.70**	29.83	30.46	28.29	**28.99**	23.58	24.53	19.80	**21.29**
Hill	**37.31**	37.16	**33.84**	33.68	**32.06**	31.89	**30.85**	30.71	**27.13**	26.99	24.10	24.30
	37.00	37.17	33.49	33.69	31.67	31.86	30.47	30.61	26.91	26.86	**24.37**	24.13
House	**40.13**	40.00	**37.06**	37.05	35.31	**35.32**	**34.03**	34.16	29.53	**29.90**	25.20	25.63
	39.04	39.91	35.81	36.80	34.21	35.11	33.08	33.97	28.91	29.63	25.44	**25.65**
Man	**37.99**	37.84	**34.18**	34.03	**32.12**	31.98	**30.73**	30.61	**26.84**	26.72	23.86	**24.00**
	37.67	37.78	33.90	33.96	31.89	31.89	30.53	30.52	26.63	26.60	23.96	**23.97**
Peppers	**38.30**	38.15	**34.94**	34.80	**33.01**	32.87	**31.61**	31.47	**26.94**	26.87	**23.05**	23.14
	37.93	38.06	34.51	34.66	32.56	32.70	31.18	31.26	26.60	26.53	22.93	22.64
Straw	35.81	**35.92**	31.46	31.39	29.13	28.95	**27.52**	27.36	**22.79**	22.67	19.42	**19.50**
	35.36	35.87	30.84	**31.50**	28.50	**29.13**	26.93	27.50	22.00	22.48	18.95	19.23
Average	**37.98**	37.87	**34.40**	34.24	**32.42**	32.23	**31.04**	30.85	**26.71**	26.54	23.10	**23.29**
	37.49	37.84	33.79	34.22	31.78	32.21	30.39	30.83	26.04	26.44	22.91	23.14

Fig. 5. Deblurring performance comparison on the *Cameraman* image. From left to right and top to bottom: noisy and blurred image (9×9 uniform blur, $\sigma_n = \sqrt{2}$), the deblurred images by IDD-BM3D [41] (PSNR = 28.56 dB; FSIM = 0.9007), ASDS-Reg [21] (PSNR = 28.08 dB; FSIM = 0.8950), and the proposed NCSR (PSNR = **28.62** dB; FSIM = **0.9026**).

Fig. 6. Deblurring performance comparison on the *Starfish* image. From left to right and top to bottom: noisy and blurred image (9×9 uniform blur, $\sigma_n = \sqrt{2}$), the deblurred images by IDD-BM3D [41] (PSNR = 29.48 dB; FSIM = 0.9167), ASDS-Reg [21] (PSNR = 29.72 dB; FSIM = 0.9208), and the proposed NCSR (PSNR = **30.28** dB; FSIM = **0.9293**).

IDD-BM3D deblurring method [41], and the adaptive sparse domain selection method (denoted by ASDS-Reg) [21]. Note that the recently proposed IDD-BM3D method is an improved version of BM3D deblurring method [20], and ASDS-Reg is a very competitive sparsity-based deblurring method with adaptive sparse domain selection.

The PSNR and FSIM results on a set of 10 photographic images are reported in Table II. From Table II, we can conclude that the proposed NCSR deblurring method outperforms much the other competing methods. In average NCSR outperforms IDD-BM3D by 0.3 dB and 0.17 dB for the uniform blur and Gaussian blur, respectively. The visual comparisons of the deblurring methods are shown in Figs. 5∼6, from which we can see that the NCSR method produces much cleaner and sharper image edges and textures than other methods.

The PSNR results for the 6 typical deblurring experiments presented in [36] and [41] are reported in Table III. For fair comparison, the PSNR results of other competing methods are directly obtained from [41]. We optimize the parameters of the proposed deblurring method for each experiment. From Table III, we can see that both the IDD-BM3D method of [41] and the proposed NCSR method can

TABLE II
PSNR (dB) AND FSIM RESULTS BY DIFFERENT DEBLURRING METHODS

	9×9 Uniform Blur, $\sigma_n = \sqrt{2}$										
Images	*Butterfly*	*Boats*	*C. Man*	*House*	*Parrot*	*Lena*	*Barbara*	*Starfish*	*Peppers*	*Leaves*	**Average**
FISTA	28.37	29.04	26.82	31.99	29.11	28.33	25.75	27.75	28.43	26.49	28.21
[36]	0.9119	0.8858	0.8627	0.9017	0.9002	0.8798	0.8375	0.8775	0.8813	0.8958	0.8834
l_0-SPAR	27.10	29.86	26.97	32.98	29.34	28.72	26.42	28.11	28.66	26.30	28.44
[37]	0.8879	0.9094	0.8689	0.9225	0.9262	0.9063	0.8691	0.8951	0.9066	0.8776	0.8970
IDD-BM3D	29.21	**31.20**	28.56	**34.44**	31.06	29.70	27.98	29.48	29.62	29.38	30.06
[42]	**0.9287**	**0.9304**	0.9007	0.9369	0.9364	0.9197	0.9014	0.9167	0.9200	0.9295	0.9220
ASDS-Reg	28.70	30.80	28.08	34.03	31.22	29.92	27.86	29.72	29.48	28.59	29.84
[21]	0.9053	0.9236	0.8950	0.9337	0.9306	**0.9256**	0.9088	0.9208	0.9203	0.9075	0.9171
NCSR	**29.68**	31.08	**28.62**	34.31	**31.95**	**29.96**	**28.10**	**30.28**	**29.66**	**29.98**	**30.36**
	0.9271	0.9294	**0.9026**	**0.9415**	**0.9411**	0.9254	**0.9117**	**0.9293**	**0.9220**	**0.9341**	**0.9263**
	Gaussian Blur, $\sigma_n = \sqrt{2}$										
FISTA	30.36	29.36	26.81	31.50	31.23	29.47	25.03	29.65	29.42	29.36	29.22
[36]	**0.9452**	0.9024	0.8845	0.8968	0.9290	0.9011	0.8415	0.9256	0.9057	0.9393	0.9071
IDD-BM3D	30.73	**31.68**	28.17	**34.08**	32.89	**31.45**	27.19	31.66	29.99	31.40	30.92
[42]	0.9442	**0.9426**	**0.9136**	**0.9359**	0.9561	**0.9430**	0.8986	0.9496	**0.9373**	**0.9512**	**0.9372**
ASDS-Reg	29.83	30.27	27.29	31.87	32.93	30.36	27.05	31.91	28.95	30.62	30.11
[21]	0.9126	0.9064	0.8637	0.8978	0.9576	0.9058	0.8881	0.9491	0.9039	0.9304	0.9115
NCSR	**30.84**	31.49	**28.34**	33.63	**33.39**	31.26	**27.91**	**32.27**	**30.16**	**31.57**	**31.09**
	0.9381	0.9371	0.9078	0.9333	**0.9587**	0.9389	**0.9088**	**0.9551**	0.9331	**0.9508**	0.9362

TABLE III
COMPARISON OF THE PSNR (dB) RESULTS OF THE DEBLURRING METHODS

	Scenario						Scenario					
	1	2	3	4	5	6	1	2	3	4	5	6
Method	*Cameraman* (256×256)						*House* (256×256)					
BSNR	31.87	25.85	40.00	18.53	29.19	17.76	29.16	23.14	40.00	15.99	26.61	15.15
Input PSNR	22.23	22.16	20.76	24.62	23.36	29.82	25.61	25.46	24.11	28.06	27.81	29.98
TVMM [7]	7.41	5.17	8.54	2.57	3.36	1.30	7.98	6.57	10.39	4.12	4.54	2.44
L0-Spar [37]	7.70	5.55	9.10	2.93	3.49	1.77	8.40	7.12	11.06	4.55	4.80	2.15
IDD-BM3D [42]	**8.85**	**7.12**	**10.45**	**3.98**	4.31	**4.89**	9.95	**8.55**	12.89	5.79	**5.74**	**7.13**
NCSR	8.78	6.69	10.33	3.78	**4.60**	4.50	**9.96**	8.48	**13.12**	**5.81**	5.67	6.94
	Lena (512×512)						*Barbara* (512×512)					
BSNR	29.89	23.87	40.00	16.47	27.18	15.52	30.81	24.79	40.00	17.35	28.07	16.59
Input PSNR	27.25	27.04	25.84	28.81	29.16	30.03	23.34	23.25	22.49	24.22	23.77	29.78
TVMM [7]	6.36	4.98	7.47	3.52	3.61	2.79	3.10	1.33	3.49	0.41	0.75	0.59
L0-Spar [37]	6.66	5.71	7.79	4.09	4.22	1.93	3.51	1.53	3.98	0.73	0.81	1.17
IDD-BM3D [42]	7.97	**6.61**	8.91	**4.97**	4.85	**6.34**	7.64	**3.96**	**6.05**	1.88	1.16	5.45
NCSR	**8.03**	6.54	**9.25**	4.93	**4.86**	6.19	**7.76**	3.64	5.92	**2.06**	**1.43**	**5.50**

achieve significant PSNR improvement over other competing methods. Parts of the deblurred *Barbara* image (scenario 4) by the competing methods are shown in Fig. 7. We can see that IDD-BM3D and the proposed NCSR method can better recover the fine textures than other competing methods. Moreover, the textures recovered by the proposed NCSR method are better than those by the IDD-BM3D method.

We also test the proposed NCSR deblurring method on real motion blurred images. Since the blur kernel estimation is a non-trivial task, we borrowed the kernel estimation method from [34] to estimate the blur kernel and apply the estimated blur kernel in NCSR to restore the original images. In Fig. 8 we present the deblurring results by the blind deblurring method of [34] and the proposed NCSR method. We can see that the images restored by our approach are much clearer and much more details are recovered. Considering that the estimated kernel will have bias from the true unknown blurring kernel, these experiments validate that NCSR is robust to the kernel estimation errors.

TABLE IV

PSNR (dB) AND FSIM RESULTS (LUMINANCE COMPONENTS) OF THE RECONSTRUCTED HR IMAGES

	Noiseless									
Images	*Butterfly*	*flower*	*Girl*	*Pathenon*	*Parrot*	*Raccoon*	*Bike*	*Hat*	*Plants*	Average
TV [26]	26.56	27.51	31.24	26.00	27.85	27.54	23.66	29.20	31.34	27.88
	0.8970	0.8295	0.8598	0.7701	0.9081	0.7950	0.8032	0.8623	0.8909	0.8462
Sparsity [23]	24.70	27.87	32.87	26.27	28.70	28.51	23.23	29.63	31.55	28.15
	0.7995	0.8611	0.8989	0.8018	0.9233	0.8490	0.8138	0.8705	0.8964	0.8571
ASDS-Reg	27.28	29.33	33.50	27.00	30.45	29.25	24.67	30.99	33.44	29.55
[21]	0.8789	0.8960	0.9117	0.8235	0.9421	0.8750	0.8566	0.9018	0.9242	0.8900
NCSR	**28.10**	**29.50**	**33.65**	**27.19**	**30.50**	**29.28**	**24.74**	**31.27**	**34.00**	**29.80**
	0.9031	**0.9004**	**0.9210**	**0.8367**	**0.9452**	**0.8796**	**0.8606**	**0.9093**	**0.9369**	**0.8992**
	Noisy									
TV [26]	25.49	26.57	29.86	25.35	27.01	26.74	23.11	28.13	29.70	26.88
	0.8707	0.8147	0.8342	0.7706	0.8562	0.7888	0.7890	0.8340	0.8561	0.8238
Sparsity [23]	23.61	26.60	30.71	25.40	27.15	27.22	22.45	28.31	29.57	26.78
	0.7764	0.8208	0.8685	0.7814	0.8632	**0.8243**	0.7807	0.8237	0.8524	0.8213
ASDS-Reg [21]	26.06	27.83	31.87	26.22	29.01	**28.01**	23.62	29.61	31.18	28.16
	0.8530	**0.8601**	**0.8731**	**0.7904**	0.9182	0.8202	**0.8200**	0.8630	0.8837	0.8535
NCSR	**26.86**	**28.08**	**32.03**	**26.38**	**29.51**	28.03	**23.80**	**29.94**	**31.73**	**28.48**
	0.8860	0.8633	0.8741	0.7969	**0.9210**	0.8113	0.8241	**0.8700**	**0.8955**	**0.8602**

Fig. 7. Deblurring performance comparison on the *Barbara* (512×512) image. From left to right and top to bottom: original image, noisy, and blurred image (scenario 4: PSF$=[1\ 4\ 6\ 4\ 1]^T[1\ 4\ 6\ 4\ 1]/256$, $\sigma_n = 7$), the deblurred images by IDD-BM3D [41] (PSNR$=26.10$ dB) and the proposed NCSR (PSNR$=$**26.28** dB).

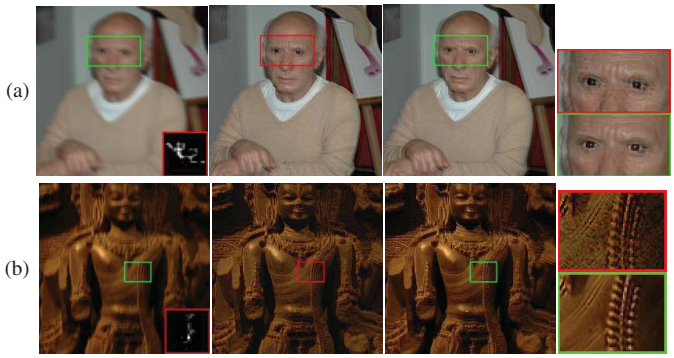

Fig. 8. Deblurring performance comparison on real motion blurred images with the blur kernel estimated using the kernel estimation approach from [34]. (a) and (b) Motion blurred image, deblurred image by [34], deblurred image by proposed NCSR, and close-up views.

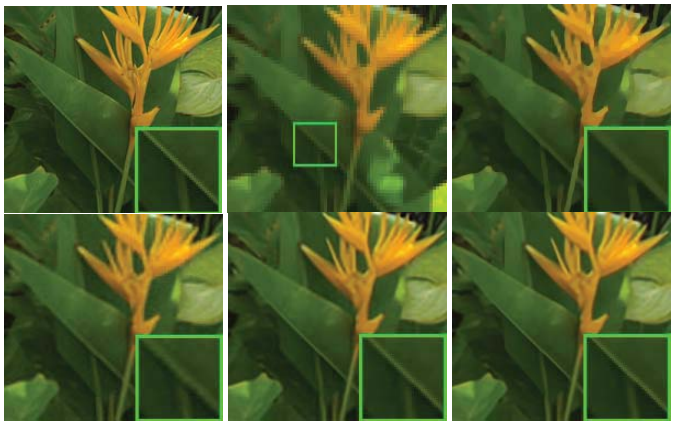

Fig. 9. Image super-resolution performance comparison on *Plant* image (scaling factor 3, $\sigma_n = 0$). From left to right and top to bottom: original image, LR image, the reconstructed images by TV [25] (PSNR$=31.34$ dB; FSIM$=0.8909$), sparsity-based [23] (PSNR$=31.55$ dB; FSIM$=0.8964$), ASDS-Reg [21] (PSNR$=33.44$ dB; FSIM$=0.9242$), and the proposed NCSR (PSNR$=$**34.00** dB; FSIM$=$**0.9369**).

C. Image Super-Resolution

In image super-resolution the simulated LR image is generated by first blurring an HR image with a 7×7 Gaussian kernel with standard deviation 1.6, and then downsampling the blurred image by a scaling factor 3 in both horizontal and vertical directions. The additive Gaussian noise of standard deviation 5 is also added to the LR images, making the IR problem more challenging. Since human visual system is more sensitive to luminance changes, we only apply the IR methods to the luminance component and use the simple bicubic interpolator for the chromatic components.

We compare the proposed NCSR approach with three recently developed image super-resolution methods, including the TV-based method [25], the sparse representation based

Fig. 10. Image super-resolution performance comparison on *Parrot* image (scaling factor 3, $\sigma_n = 5$). From left to right and top to bottom: original image, LR image, the reconstructed images by TV [25] (PSNR = 27.01 dB; FSIM = 0.8562), sparsity-based [23] (PSNR = 27.15 dB; FSIM = 0.8632), ASDS-Reg [21] (PSNR = 29.01 dB; FSIM = 0.9182), and the proposed NCSR (PSNR = **29.51** dB; FSIM = **0.9210**).

method [23], and the ASDS-Reg method [21].[1] Since the sparsity-based method in [23] cannot perform the resolution upscaling and deblurring simultaneously, as suggested by the authors [23] we apply the iterative back-projection [37] to the output of method [23] to remove the blur.

The PSNR results of the test methods on a set of 9 natural images are reported in Table IV, from which we can conclude that the proposed NCSR approach significantly outperforms the TV [25] and sparsity-based methods [23]. The subjective comparison between the NCSR and other methods are shown in Figs. 9∼10. Obviously, the NCSR approach reconstruct the best visually pleasant HR images. The reconstructed edges are much sharper than all the other three competing methods, and more image fine structures are recovered.

V. Conclusion

In this paper we presented a novel nonlocally centralized sparse representation (NCSR) model for image restoration. The sparse coding noise (SCN), which is defined as the difference between the sparse code of the degraded image and the sparse code of the unknown original image, should be minimized to improve the performance of sparsity-based image restoration. To this end, we proposed a centralized sparse constraint, which exploits the image nonlocal redundancy, to reduce the SCN. The Bayesian interpretation of the NCSR model was provided and this endows the NCSR model an iteratively reweighted implementation. An efficient iterative shrinkage function was presented for solving the l_1-regularized NCSR minimization problem. Experimental results on image denoising, deblurring and super-resolution demonstrated that the NCSR approach can achieve highly competitive performance to other leading denoising methods, and outperform much other leading image deblurring and super-resolution methods.

[1]We thank the authors of [18], [21], [23], [33]–[36], [39]–[41] for providing their source codes or experimental results. The code associated with this paper will be available online.

References

[1] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[2] D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[3] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[4] M. Bertero and P. Boccacci, *Introduction to Inverse Problems Imaging*. Bristol, U.K.: IOP Publishing, 1998.

[5] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, Nov. 1992.

[6] T. Chan, S. Esedoglu, F. Park, and A. Yip, "Recent developments in total variation image restoration," in *Mathematical Models of Computer Vision*, N. Paragios, Y. Chen, and O. Faugeras, Eds. New York: Springer Verlag, 2005.

[7] J. Oliveira, J. M. Bioucas-Dias, and M. Figueiredo, "Adaptive total variation image deblurring: A majorization-minimization approach," *Signal Process.*, vol. 89, no. 9, pp. 1683–1693, Sep. 2009.

[8] A. N. Tikhonov, "Solution of incorrectly formulated problems and regularization method," *Soviet Math. Dokl.*, vol. 4, no. 4, pp. 1035–1038, 1963.

[9] I. Daubechies, M. Defriese, and C. DeMol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.

[10] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Soc. Ind. Appl. Math. J. Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, 2005.

[11] 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.

[12] 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.

[13] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2992–3004, Dec. 2007.

[14] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *Soc. Ind. Appl. Math. Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[15] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, Jun. 2010.

[16] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[17] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[18] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, Tokyo, Japan, Sep.–Oct. 2009, pp. 2272–2279.

[19] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[20] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," *Proc. SPIE*, vol. 6812, pp. 681207-1–681207-12, Jul. 2008.

[21] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1838–1857, Jul. 2011.

[22] W. Dong, L. Zhang, and G. Shi, "Centralized sparse representation for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Nov. 2011, pp. 1259–1266.

[23] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[24] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *Soc. Ind. Appl. Math. J. Imaging Sci.*, vol. 3, no. 3, pp. 253–276, 2010.

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

[26] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[27] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. SPIE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.
[28] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.
[29] I. Ramirez and G. Sapiro, "Universal regularizers for robust sparse coding and modeling," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3850–3964, Sep. 2012.
[30] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.
[31] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Tokyo, Japan, Sep.–Oct. 2009, pp. 349–356.
[32] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
[33] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 479–486.
[34] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single image," *ACM Trans. Graph. (SIGGRAPH)*, vol. 25, no. 3, pp. 787–794, 2006.
[35] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.
[36] J. Portilla, "Image restoration through l0 analysis-based sparse optimization in tight frames," in *Proc. IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 3909–3912.
[37] M. Irani and S. Peleg, "Motion analysis for image enhancement: Resolution, occlusion, and transparency," *J. Visual Commun. Image Represent.*, vol. 4, no. 4, pp. 324–335, Dec. 1993.
[38] E. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.
[39] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising," *Int. J. Comput. Vis.*, vol. 86, no. 1, pp. 1–32, Jan. 2010.
[40] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *Proc. Workshop Signal Process. Adaptive Sparse Struct. Represent.*, Saint-Malo, France, Apr. 2009, pp. 1–6.
[41] A. Danielyan, V. Katkovnik, and K. Egiazarian, "BM3D frames and variational image deblurring," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1715–1728, Apr. 2012.
[42] M. Elad and I. Yavneh, "A plurality of sparse representation is better than the sparsest one alone," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4701–4714, Oct. 2009.

Weisheng Dong received the B.S. degree in electronic engineering from the Huazhong University of Science and Technology, Wu Han, China, and the Ph. D. degree in circuits and system from the Xidian University, Xi'an, China, in 2004 and 2010, respectively. He was a Visiting Student with Microsoft Research Asia, Beijing, China, in 2006.

He was a Research Assistant with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong, from 2009 to 2010. In 2010, he joined Xidian University as a Lecturer, where he has been an Associate Professor since 2012. His current research interests include inverse problems in image processing, sparse signal representation, and image compression.

Dr. Dong was a recipient of the Best Paper Award at the SPIE Visual Communication and Image Processing 2010.

Lei Zhang (M'04) received the B.S. degree from Shenyang Institute of Aeronautical Engineering, Shenyang, China, in 1995, and the M.S. and Ph.D. degrees in automatic control theory and engineering from Northwestern Polytechnical University, Xi'an, China, in 1998 and 2001, respectively.

He was a Research Associate with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong, from 2001 to 2002. From 2003 to 2006, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada. In 2006, he joined the Department of Computing, The Hong Kong Polytechnic University, as an Assistant Professor, where he has been an Associate Professor since 2010. His current research interests include image and video processing, computer vision, pattern recognition, and biometrics.

Dr. Zhang was a recipient the Faculty Merit Award for research and scholarly activities in 2010 and 2012, the Most Valued Paper published in *Pattern Recognition Journal* in 2010, and the Best Paper Award at SPIE VCIP2010. He is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS and Reviews and *Image and Vision Computing Journal*.

Guangming Shi (SM'10) received the B.S. degree in automatic control, the M.S. degree in computer control, and the Ph.D. degree in electronic information technology from Xidian University, Xi'an, China, in 1985, 1988 and 2002, respectively. He had studied with the Department of Electronic Engineering at University of Illinois at Urbana-Champaign, Champaign, in 2004.

He joined the School of Electronic Engineering, Xidian University, in 1988. From 1994 to 1996, he was a Research Assistant with the Department of Electronic Engineering, University of Hong Kong. Since 2003, he has been a Professor with the School of Electronic Engineering, Xidian University. In 2004, he was the Head of National Instruction Base of Electrician & Electronic (NIBEE). He is currently the Deputy Director of the School of Electronic Engineering, Xidian University, and the Academic Leader on circuits and systems. He has authored or co-authored over 60 papers in journals and conferences. His current research interests include compressed sensing, theory and design of multirate filter banks, image denoising, low-bit-rate image and video coding, and implementation of algorithms for intelligent signal processing (using digital signal processing and field-programmable gate arrays).

Xin Li (M'00–SM'08) received the B.S. degree (highest Hons.) in electronic engineering and information science from the University of Science and Technology of China, Hefei, China, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1996 and 2000, respectively.

He was a Member of Technical Staff with Sharp Laboratories of America, Camas, WA, from 2000 to 2002. Since 2003, he has been a Faculty Member with Lane Department of Computer Science and Electrical Engineering. His current research interests include image and video coding and processing.

Dr. Li was a recipient of the Best Student Paper Award at the Conference of Visual Communications and Image Processing as the junior author in 2001, the Runner-Up Prize of Best Student Paper Award at the IEEE Asilomar Conference on Signals, Systems, and Computers as the senior author in 2006, and the Best Paper Award at the Conference of Visual Communications and Image Processing as the single author in 2010. He is currently a member of the Image, Video, and Multidimensional Signal Processing Technical Committee and an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING.