Image Compressive Sensing Recovery via Collaborative Sparsity

Jian Zhang, *Student Member, IEEE*, Debin Zhao, *Member, IEEE*, Chen Zhao, Ruiqin Xiong, *Member, IEEE*, Siwei Ma, *Member, IEEE*, and Wen Gao, *Fellow, IEEE*

***Abstract*—Compressive sensing (CS) has drawn quite an amount of attention as a joint sampling and compression approach. Its theory shows that when the signal is sparse enough in some domain, it can be decoded from many fewer measurements than suggested by the Nyquist sampling theory. So one of the most challenging researches in CS is to seek a domain where a signal can exhibit a high degree of sparsity and hence be recovered faithfully. Most of the conventional CS recovery approaches, however, exploited a set of fixed bases (e.g., DCT, wavelet, and gradient domain) for the entirety of a signal, which are irrespective of the nonstationarity of natural signals and cannot achieve high enough degree of sparsity, thus resulting in poor rate-distortion performance. In this paper, we propose a new framework for image compressive sensing recovery via collaborative sparsity, which enforces local 2-D sparsity and nonlocal 3-D sparsity simultaneously in an adaptive hybrid space-transform domain, thus substantially utilizing intrinsic sparsity of natural images and greatly confining the CS solution space. In addition, an efficient augmented Lagrangian-based technique is developed to solve the above optimization problem. Experimental results on a wide range of natural images are presented to demonstrate the efficacy of the new CS recovery strategy.**

***Index Terms*—Augmented Lagrangian, compressive sensing (CS), image recovery, sparsity.**

I. Introduction

THE RECENT development of compressive sensing (CS) theory has drawn quite an amount of attention as an alternative to the current methodology of sampling followed by compression [1]–[3]. By exploiting the redundancy existed in a signal, CS conducts sampling and compression at the same time. From many fewer acquired measurements than suggested by the Nyquist sampling theory, CS theory demonstrates that, a signal can be reconstructed with high probability when it exhibits sparsity in some domain.

Manuscript received March 01, 2012; revised July 04, 2012; accepted September 11, 2012. Date of publication October 26, 2012; date of current version December 05, 2012. This work was supported in part by the Major State Basic Research Development Program of China (2009CB320905) and in part by the National Science Foundation (61073083 and 61121002). This paper was recommended by Guest Editor R. Rovatti.

J. Zhang and D. Zhao are with the School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China (e-mail: jzhangcs@hit.edu.cn; dbzhao@hit.edu.cn).

C. Zhao, R. Xiong, S. Ma, and W. Gao are with the National Engineering Laboratory for Video Technology, and Key Laboratory of Machine Perception (MoE), School of Electrical Engineering and Computer Science, Peking University, Beijing 100871, China (e-mail: zhaochen@pku.edu.cn; rqxiong@pku.edu.cn; swma@pku.edu.cn; wgao@pku.edu.cn).

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/JETCAS.2012.2220391

In CS theory, a signal is usually sampled by a random projection that is signal-independent and reconstructed by minimizing the ℓ_0 or ℓ_1 optimization problem with the prior that the signal is sparse in some transformation domain. Since the ℓ_0 minimization is discontinuous and an NP-hard problem, the most common one is to use the ℓ_1 norm, which is the optimal convex approximation of ℓ_0 norm and has been proved that for many problems it is probable that the ℓ_1 norm is equivalent to the ℓ_0 norm in a technical sense. This equivalence result allows one to solve the ℓ_1 problem, which is easier than the ℓ_0 problem. Many CS recovery algorithms are recently proposed: linear programming [4], gradient projection sparse reconstruction [5], matching pursuit [6], and iterative thresholding [7].

However, in order to reconstruct the signal exactly and efficiently, the convex optimization method with ℓ_1 norm often needs more observations than directly minimizing the ℓ_0 norm. Some algorithms are then proposed to reconstruct the original signal with fewer measurements than those needed by the ℓ_0 norm minimization method. A nonconvex method via solving a series of weighted ℓ_1 minimization problems was proposed in [8], where the weight for the entries with a larger value is iteratively decreased. Due to the fact that the ℓ_0 norm is a discontinuous, Mohimani *et al.* [9] proposed to approximate it by a series of suitable continuous ones and minimize it by means of a minimization algorithm for continuous functions. Babacan *et al.* [10] utilized a hierarchical form of the Laplace prior to model the sparsity of the unknown signal. Wang *et al.* [11] developed a sparse signal reconstruction method by an iterative support detection technique, aiming to reduce the requirement on the number of measurements. Alternatively, Baron *et al.* [12] performed approximate Bayesian inference using belief propagation (BP) decoding, which represents the CS encoding matrix as a graphical model. All these algorithms try to solve the combinational optimization problem more efficiently with the tool of nonconvex optimization method.

An attractive strength of CS-based compression is that the encoder is made signal-independent and computationally inexpensive at the cost of high decoder complexity, namely, simple encoder and complex decoder, which quite resembles distributed source coding in spirit. More specifically, in the encoding process, the same random projection can be conducted on all input signals, which is nonadaptive and irrespective of any differences in their structures. It is up to the decoder to solve a large scale optimization problem to recover the randomly sampled signal in a domain where the signal exhibits sparsity. This asymmetric design is severely desirable in some image processing applications when the data acquisition devices must

be simple (e.g., inexpensive resource-deprived sensors), or when oversampling can harm the object being captured (e.g., X-ray imaging) [13].

CS theory shows that the sparsity degree of a signal plays a significant role in recovery. The higher degree of a signal, the higher recovery quality it will have. So, seeking a domain in which the signal has a high degree of sparsity is one of the main challenges CS recovery should face. However, natural signals such as images are typically nonstationary, there exists no universal domain in which all parts of the signals are sparse. The most current CS recovery methods explore a set of fixed domains (e.g., DCT, wavelet, and gradient domain) [14]–[16], and therefore are signal-independent or not adaptive, resulting in poor rate-distortion performance compared to the conventional coding techniques.

Towards this problem, Wu *et al.* [13] proposed a model-guided adaptive recovery of compressive sensing (MARX) utilizing a piecewise autoregressive model to adapt to the changing second order statistics of natural images. Many recent works incorporated additional prior knowledge about transform coefficients (statistical dependencies, structure, etc.) into the CS recovery framework, such as Gaussian scale mixtures (GSM) models [17], tree-structured wavelet (TSW) [18], tree-structured DCT (TSDCT) [19]. Additionally, in [20], a projection-driven CS recovery coupled with block-based random image sampling is developed, which aims to encourage sparsity in the domain of directional transforms.

Considering the fact that the natural image signal is nonstationary and its sparse domain varies spatially, in this paper, we first establish a new sparsity measure, called collaborative sparsity measure (CoSM), and then propose a novel strategy for CS recovery via collaborative sparsity (RCoS). Part of our previous work has been published in [33]. The collaborative sparsity enforces local 2-D sparsity and nonlocal 3-D sparsity simultaneously, which offers a powerful mechanism of characterizing the structured sparsity of a natural image and enables a natural image to be highly sparse in an adaptive hybrid space-transform domain. To make RCoS tractable and robust, an augmented Lagrangian-based technique is developed to efficiently solve the above severely underdetermined inverse problem. Extensive experiments on a wide range of CS-acquired images manifest that RCoS is able to increase recovery quality by a large margin compared with the conventional CS recovery methods or require many fewer measurements for a given reconstruction quality.

The remainder of the paper is organized as follows. Section II briefly reviews CS theory and introduces the classic augmented Lagrangian method for constrained optimization problem. Section III provides the design of collaborative sparsity measure (CoSM) in details. Section IV shows how CoSM is incorporated into the framework of CS recovery and gives the implementation details of RCoS. Experimental results are reported in Section V. In Section VI, we conclude this paper.

II. Background

A. Compressive Sensing

A signal $\boldsymbol{u}$ of size N is said to be sparse in domain Ψ, if its transform coefficients are mostly zeros, or nearly sparse if the dominant portion of coefficients are either zeros or very close to zeros. The sparsity of $\boldsymbol{u}$ in Ψ is quantified by the number of significant elements within the coefficients vector $\Psi\boldsymbol{u}$.

More specifically, given M linear measurements, the CS recovery of $\boldsymbol{u}$ from $\boldsymbol{b}$ is formulated as the following constrained optimization problem:

$$\min_{u} \|\Psi\boldsymbol{u}\|_p \quad \text{s.t.} \quad \boldsymbol{b} = \boldsymbol{A}\boldsymbol{u} \tag{1}$$

where $\boldsymbol{A}$ represents the random projections (RS). p is usually set to 1 or 0, characterizing the sparsity of the vector $\Psi\boldsymbol{u}$. $||*||_1$ is ℓ_1 norm, adding all the absolute values of the entries in a vector, while $||*||_0$ is ℓ_0 norm, counting the nonzero entries of a vector. According to [21], CS is capable of recovering K-sparse signal $\boldsymbol{u}$ (with an overwhelming probability) from $\boldsymbol{b}$ of size M, provided that the number of random samples meets $M \geq \alpha K(N/K)$. The required sampling rate (M/K), to incur lossless recovery, is roughly proportional to (K/N). A compressive imaging camera prototype using RS has been presented in [22].

B. Augmented Lagrangian Method

Consider the constrained optimization problem

$$\min_{\boldsymbol{f}} \boldsymbol{E}(\boldsymbol{f}) \quad \text{s.t.} \quad \boldsymbol{H}\boldsymbol{f} = \boldsymbol{g} \tag{2}$$

where $\boldsymbol{f} \in \mathbb{R}^n$, $\boldsymbol{g} \in \mathbb{R}^m$, and $\boldsymbol{H} \in \mathbb{R}^{m\times n}$, i.e., there are m linear equality constraints. The so-called augmented Lagrangian function for this problem is defined as

$$\mathcal{L}_{\mathcal{A}}(\boldsymbol{f}, \boldsymbol{\lambda}) = \boldsymbol{E}(\boldsymbol{f}) - \boldsymbol{\lambda}^T(\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}) + \frac{\mu}{2}\|\boldsymbol{H}\boldsymbol{f} - \boldsymbol{g}\|_2^2 \tag{3}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^m$ is a vector of Lagrangian multipliers and $\mu \geq 0$ is called the penalty parameter [23].

The basic idea of the *augmented Lagrangian method* (ALM) is to find a saddle point of $\mathcal{L}_{\mathcal{A}}(\boldsymbol{f}, \boldsymbol{\lambda})$, which is also the solution of the original (2). ALM consists in minimizing $\mathcal{L}_{\mathcal{A}}(\boldsymbol{f}, \boldsymbol{\lambda})$ with respect to $\boldsymbol{f}$, keeping $\boldsymbol{\lambda}$ fixed, then updating $\boldsymbol{\lambda}$, and repeating these two steps until some convergence criterion is satisfied. Formally, the ALM works as follows [24]:

Algorithm *ALM*

1. Set $k = 0$, choose $\mu > 0$, $\boldsymbol{f}_0$ and $\boldsymbol{\lambda}_0$.
2. **repeat**
3. $\boldsymbol{f}_{k+1} \in \operatorname{argmin}_{\boldsymbol{f}} \mathcal{L}_{\mathcal{A}}(\boldsymbol{f}, \boldsymbol{\lambda}_k)$
4. $\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k - \mu\left(\boldsymbol{H}\boldsymbol{f}_{k+1} - \boldsymbol{g}\right)$
5. $k \leftarrow k + 1$
6. **until** stopping criterion is satisfied.

The augmented Lagrangian function differs from the standard Lagrangian function by adding a square penalty term, and differs from the quadratic penalty function by the presence of the linear term involving the multiplier $\boldsymbol{\lambda}$. In this respect, the augmented Lagrangian function is a combination of the Lagrangian and quadratic penalty functions [16].

In order to facilitate the discussions in the following optimization section, we briefly introduce two lemmas [16], [24].

Lemma 1: The minimization problem

$$\mathbf{argmin}_{\boldsymbol{x}} = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{a}\|_2^2 + \tau \cdot \|\boldsymbol{x}\|_1$$

has a closed form, which can be expressed as

$$\hat{\boldsymbol{x}} = \mathbf{soft}(\boldsymbol{a}, \tau) = \mathbf{sgn}(\boldsymbol{a}, \tau) \cdot \mathbf{max}\left(\mathbf{abs}(\boldsymbol{a}) - \tau, 0\right).$$

Lemma 2: The minimization problem

$$\mathbf{argmin}_{\boldsymbol{x}} = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{a}\|_2^2 + \tau \cdot \|\boldsymbol{x}\|_0$$

has a closed form, which can be expressed as

$$\hat{\boldsymbol{x}} = \mathbf{hard}(\boldsymbol{a}, \sqrt{2\tau}) = \boldsymbol{a} \cdot \mathbf{1}(\mathbf{abs}(\boldsymbol{a}) - \sqrt{2\tau}).$$

Here, $\cdot$ stands for the element-wise product of two vectors.

III. Collaborative Sparsity Measure in Hybrid Space-Transform Domain

As mentioned before, on one hand, one of the most significant challenges in CS is to seek a domain where a signal can be represented sparsely and hence be recovered faithfully. On the other hand, CS recovery is typically an image linear inverse problem, for which it has been demonstrated that image priors play a key role to achieve high-quality results.

Integrating the above two points, one straightforward solution is to design a new type of sparsity measure according to image priors, which means that a natural image signal can be mapped into an adaptive domain where the image signal exhibits a high degree of this type of sparsity. This is just the motivation of this paper.

This paper utilizes two kinds of image priors, namely local smoothness and nonlocal self-similarity. The former type describes the piecewise smoothness within local region, while the latter one depicts the repetitiveness of the textures or structures in natural images within nonlocal region. Therefore, a new type of sparsity should be consistent with the two properties of natural images, which is the purpose of this work.

In this section, a generic sparsity measure, called collaborative sparsity measure (CoSM) for high fidelity image CS recovery is established in a data-adaptive hybrid space-transform domain by merging two complementary sparsities—local 2-D sparsity in space domain $\Psi_{\boldsymbol{L2D}}$ and nonlocal 3-D sparsity in transform domain $\Psi_{\boldsymbol{N3D}}$. That is

$$\mathrm{CoSM}(\boldsymbol{u}) = \|\Psi_{\boldsymbol{L2D}}\boldsymbol{u}\|_p + \alpha\|\Psi_{\boldsymbol{N3D}}\boldsymbol{u}\|_q \tag{4}$$

where p and q are usually set to values from the interval [0, 1], and α is a regularization parameter, which controls the trade-off between two competing sparsity terms. $\Psi_{\boldsymbol{L2D}}$ corresponds to the above local smoothness prior and keeps image local consistency, suppressing noise effectively, while $\Psi_{\boldsymbol{N3D}}$ corresponds to the above nonlocal self-similarity prior and maintains image nonlocal consistency, retaining the sharpness and edges effectually. More details on how to design CoSM to characterize the two above properties are provided below.

A. Local 2-D Sparsity in Space Domain

Local smoothness describes the similarity of neighboring pixels in the space domain of images. That means the intensities of the neighboring pixels are quite similar.

To characterize the smoothness of images, there exist many models. Total variation (TV) model that favors the piecewise smoothness is one of the most popular regularizers and has been widely employed for image recovery in recent years [24]. From the view of statistics, a natural image is preferred when its responses for a set of filters are as small as possible [25]. That means the filtered image after some convolution with a high-passing filter is sparse (most pixels intensities are near zero). In this paper, this type of sparsity can be called local 2-D sparsity in space domain.

The widely-used filters in practice are vertical and horizontal finite difference operators, denoted by $\mathcal{D}_{\mathrm{v}}$ and $\mathcal{D}_{\mathrm{h}}$, which correspond to vertical gradient picture and horizontal gradient picture, respectively. In literatures, the gradient picture is widely modeled by generalized Gaussian distribution (GGD) [26]. Without loss of generality, in this paper, let $\Psi_{\boldsymbol{L2D}} = [\mathcal{D}] = [\mathcal{D}_{\mathrm{v}}; \mathcal{D}_{\mathrm{h}}]$ and set p to be 1 in (4) to achieve the local 2-D sparsity in space domain $\Psi_{\boldsymbol{L2D}}$, expressed as

$$\|\Psi_{\boldsymbol{L2D}}\boldsymbol{u}\|_1 = \|\mathcal{D}\boldsymbol{u}\|_1 = \|\mathcal{D}_{\mathrm{v}}\boldsymbol{u}\|_1 + \|\mathcal{D}_{\mathrm{h}}\boldsymbol{u}\|_1 \tag{5}$$

which essentially underlies the fact that an image exhibits Laplacian sparsity.

Note that $\Psi_{\boldsymbol{L2D}}$ has the same expression as anisotropic total variation [9], and can be regarded as a statistical interpretation of anisotropic TV.

Recently, many algorithms to solve CS problems with anisotropic TV regularization have been proposed [15], [16], [36]. To the best of our knowledge, [16] and [36] achieve the state-of-the-art CS recovery results. In this paper, as part of collaborative sparsity measure, $\Psi_{\boldsymbol{L2D}}$ is used to characterize the local smoothness prior of natural images. It is worth stressing that $\Psi_{\boldsymbol{L2D}}$ can also be substituted by the second order derivatives or the more sophisticated learned filters [27] and by setting p to be 1/2 or 2/3 as hyper-Laplacian priors [25]. MARX [13] also belongs to this case by utilizing auto-regressive model with adaptive filtering coefficients adjusted by the structures of images.

B. Nonlocal 3-D Sparsity in Transform Domain

Besides local smoothness, nonlocal self-similarity is another significant property of natural images, which was first proposed in the classic work for image denoising [28]. It characterizes the repetitiveness of the textures or structures embodied by natural images within nonlocal area, which can be used for retaining the sharpness and edges effectually to maintain image nonlocal consistency [28], [29], [35].

To resolve the problem of CS recovery, nonlocal self-similarity should be characterized in term of some sparsity measure. Inspired by the success of sparse representation [31] and

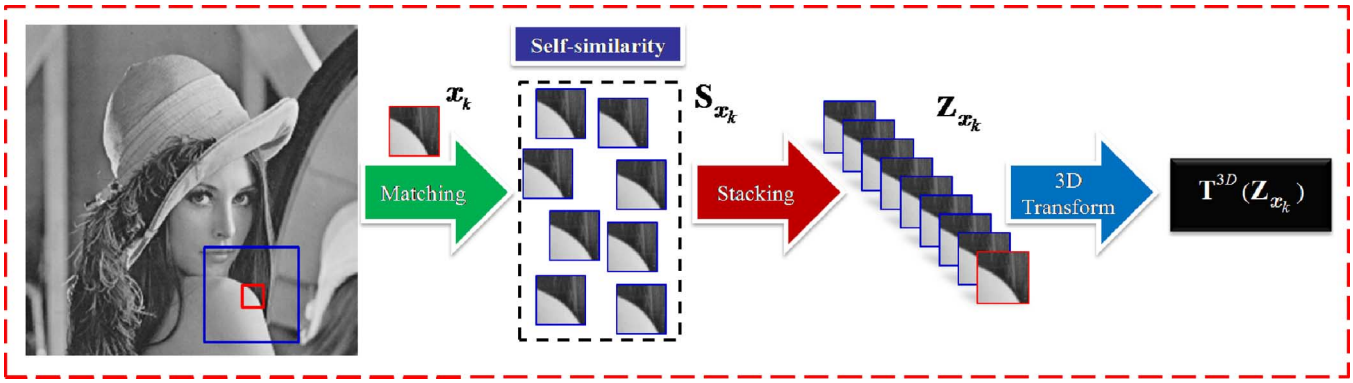


Fig. 1. Illustrations of nonlocal 3-D sparsity in transform domain.

TABLE I
DEFINITIONS ON NONLOCAL 3-D SPARSITY IN TRANSFORM DOMAIN

1: Divide the image $\boldsymbol{x}$ with size N into n overlapped blocks of size B_s and each block is denoted by $\boldsymbol{x}_k$, i.e., $k=1,2,...,n$.

2: Define $\boldsymbol{S}_{x_k}$ the set including the c best matched blocks to $\boldsymbol{x}_k$ in the $L\times L$ training window, that is, $\boldsymbol{S}_{x_k}=\left\{\boldsymbol{S}_{x_k\otimes 1},\boldsymbol{S}_{x_k\otimes 2},...,\boldsymbol{S}_{x_k\otimes c}\right\}$.

3: For every $\boldsymbol{S}_{x_k}$, a group is formed by stacking the blocks belonging to $\boldsymbol{S}_{x_k}$ into a three-dimensional array, which is denoted by Z_{x_k}.

4: Denote T^{3D} the operator of a three-dimensional transform, and $T^{3D}(Z_{x_k})$ the transform coefficients for Z_{x_k} in domain Ψ_{N3D}. Let Θ_x be the column vector with size $K=B_s*c*n$ built from all the $T^{3D}(Z_{x_k})$ arranged in lexicographic order.

self-similarity [28] in image restoration [30], we integrate them and characterize the self-similarity by means of the sparsity of the coefficients, which are achieved by transforming the 3-D group generated by stacking similar image patches. This kind of sparsity can be named as nonlocal 3-D sparsity in transform domain $\Psi_{\boldsymbol{N3D}}$.

Specifically, as illustrated in Fig. 1, for each block, we first find some blocks that are similar to it within a searching window. Then, these blocks are stacked into a 3-D array, which we call a group. Next, a 3-D transform is conducted on the 3-D array to obtain the coefficients. Finally, the number of the nonzero coefficients is used to measure the nonlocal 3-D sparsity of this patch. The nonlocal 3-D sparsity of the whole image is achieved by summing all the ones of each block.

To solve the problem, the detailed mathematical description for nonlocal 3-D sparsity in transform domain is provided in Table I.

Therefore, the mathematical formulation of the nonlocal 3-D sparsity of the whole image in transform domain $\Psi_{\boldsymbol{N3D}}$ is written as

$$\|\Psi_{\boldsymbol{N3D}}\boldsymbol{x}\|_0=\|\boldsymbol{\Theta}_{\boldsymbol{x}}\|_0=\sum_{k=1}^{n}\left\|\boldsymbol{T}^{\boldsymbol{3D}}(\boldsymbol{Z}_{\boldsymbol{x}_k})\right\|_0. \tag{6}$$

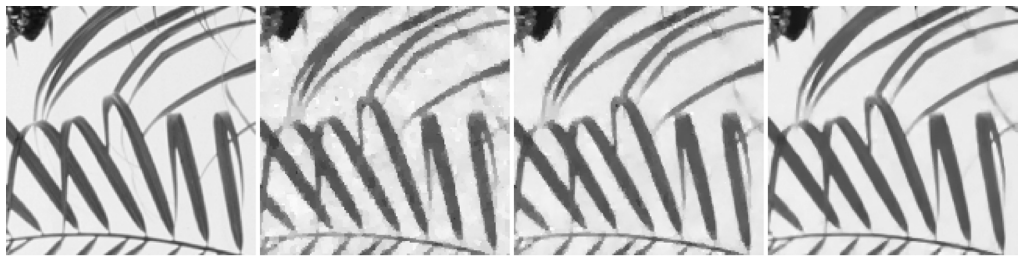
Similarly, for the convenience of computing, the inverse operator $\Omega_{\boldsymbol{N3D}}$ corresponding to $\Psi_{\boldsymbol{N3D}}$ can be defined in the following procedures. After obtaining $\boldsymbol{\Theta}_{\boldsymbol{x}}$, split it into n groups of 3-D transform coefficients, which are then inverted to generate estimates for each block in the group. The block-wise estimates are returned to their original positions and the final image estimate is achieved by averaging all of the above block-wise estimates. Therefore, given $\boldsymbol{\Theta}_{\boldsymbol{x}}$, the new estimate for $\boldsymbol{x}$ is expressed as $\hat{\boldsymbol{x}}=\Omega_{\boldsymbol{N3D}}\boldsymbol{\Theta}_{\boldsymbol{x}}$.

Here, we make a discussion about the relationship between the proposed nonlocal 3-D sparsity and the traditional nonlocal means (NLM) [28]. It is obvious that both the proposed nonlocal 3-D sparsity and NLM essentially take advantage of the property of self-similarity exhibited by natural images. The main difference is the traditional NLM for image denoising makes use of the weighted filtering thinking of the degree of similarity among similar blocks, while the proposed method characterizes self-similarity of natural images by means of the sparsity of the coefficients, which are achieved by transforming the 3-D group generated by stacking similar image blocks. The advantage of the nonlocal 3-D sparsity is that images are mapped into a high-dimensional transform space, where they exhibit high degree of sparsity, while well characterizing the self-similarity. Experiments demonstrate that this sparsity can not only reserve the common textures and details among all similar patches, but also keep the distinguished features of each block in a certain degree.

Fig. 2 shows a group of comparison results to demonstrate the effectiveness of the proposed nonlocal 3-D sparsity over the traditional NLM method, Fig. 2(a) is the original image. Fig. 2(b) is the CS recovery result by TV [16], which is regarded as local 2-D sparsity in our paper. Since the traditional NLM algorithm can be represented by the form of regularization term [29], [35], Fig. 2(c) is achieved by TV+NLM. Fig. 2(d) is the CS recovery results by local 2-D sparsity and nonlocal 3-D sparsity together, i.e., TV+ Nonlocal 3-D sparsity.

It can be observed that Fig. 2(c) is better than Fig. 2(b) in suppressing noise and preserving edge owing to the NLM regularization term in CS recovery. However, the best image quality of recovery result is obtained by Fig. 2(d) with almost invisible noise and artifacts, which fully substantiate the superiority of the proposed nonlocal 3-D sparsity over NLM.

Note that the nonlocal 3-D sparsity is data-adaptive because of its search for similar patches within nonlocal region. It is worth stressing that although (6) seems complicated as one regularization term in minimization function, we will give an effective solution by a reasonable assumption in next section.



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

Fig. 2. CS recovered results for crops of image *leaves* in the case of ratio $= 20\%$. (a) Original image. (b) CS recovery result by TV, i.e., local 2-D sparsity (PSNR $= 25.06$ dB). (c) CS recovery results by TV+NLM (PSNR $= 27.58$ dB). (d) CS recovery result by TV+Nonlocal 3-D sparsity (PSNR $= 31.33$ dB).

C. Collaborative Sparsity Measure (CoSM)

Considering local smoothness and nonlocal self-similarity in a whole, a new sparsity measure, called collaborative sparsity measure (CoSM), can be defined by combining the local 2-D sparsity in space domain and nonlocal 3-D sparsity in transform domain, which is expressed as

$$\mathrm{CoSM}(\boldsymbol{u}) = \|\Psi_{\boldsymbol{L2D}}\boldsymbol{u}\|_1 + \alpha\|\Psi_{\boldsymbol{N3D}}\boldsymbol{u}\|_0 = \|\mathcal{D}\boldsymbol{u}\|_1 + \alpha\|\boldsymbol{\Theta}_{\boldsymbol{u}}\|_0. \tag{7}$$

Note that CoSM richly characterizes local smoothness and nonlocal self-similarity of natural images, mapping a natural image into a high-dimensional data-adaptive hybrid space-transform domain $\Psi_{\boldsymbol{HST}}$ with high degree of sparsity, which greatly confines the space of CS solution.

Experimental results demonstrate that the proposed CoSM can combine the best of the both worlds, and significantly improve the reconstruction quality of CS-acquired images.

IV. CS Recovery via Collaborative Sparsity

Incorporating (5) and (6) into (7), the proposed constrained optimization problem for CS recovery via collaborative sparsity (RCoS) in an adaptive hybrid space-transform domain is formulated as

$$\min_{\boldsymbol{u}} \|\mathcal{D}\boldsymbol{u}\|_1 + \alpha\|\boldsymbol{\Theta}_{\boldsymbol{u}}\|_0 \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{u} = \boldsymbol{b}. \tag{8}$$

Note that (8) is essentially nonconvex and quite difficult to solve directly due to the nondifferentiability and nonlinearity of the collaborative sparsity term. Solving it efficiently is one of the main contributions of this paper. In this section, the implementation details of RCoS are provided.

Instead of solving (8) directly, (8) is first transformed into an equivalent variant by introducing auxiliary variables $\boldsymbol{w}$ and $\boldsymbol{x}$

$$\min_{\boldsymbol{w},\boldsymbol{u},\boldsymbol{x}} \|\boldsymbol{w}\|_1 + \alpha\|\boldsymbol{\Theta}_{\boldsymbol{x}}\|_0 \quad \text{s.t.} \quad \mathcal{D}\boldsymbol{u} = \boldsymbol{w}, \boldsymbol{u} = \boldsymbol{x}, \boldsymbol{A}\boldsymbol{u} = \boldsymbol{b}. \tag{9}$$

According to (3), the corresponding augmented Lagrangian function of (9) is

$$\begin{aligned}\mathcal{L}_{\mathcal{A}}(\boldsymbol{w},\boldsymbol{u},\boldsymbol{x}) = {} & \|\boldsymbol{w}\|_1 - \boldsymbol{\nu}^T(\mathcal{D}\boldsymbol{u} - \boldsymbol{w}) \\ & + \frac{\beta}{2}\|\mathcal{D}\boldsymbol{u} - \boldsymbol{w}\|_2^2 + \alpha\|\boldsymbol{\Theta}_{\boldsymbol{x}}\|_0 - \boldsymbol{\gamma}^T(\boldsymbol{u} - \boldsymbol{x}) \\ & + \frac{\theta}{2}\|\boldsymbol{u} - \boldsymbol{x}\|_2^2 + \frac{\mu}{2}\|\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b}\|_2^2 - \boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b})\end{aligned} \tag{10}$$

where β, θ, μ are regularization parameters associated with quadratic penalty terms $\|\mathcal{D}\boldsymbol{u} - \boldsymbol{w}\|_2^2$, $\|\boldsymbol{u} - \boldsymbol{x}\|_2^2$, $\|\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b}\|_2^2$, respectively.

The basic idea of the augmented Lagrangian method is to seek a saddle point of $\mathcal{L}_{\mathcal{A}}(\boldsymbol{w},\boldsymbol{u},\boldsymbol{x})$, which is also the solution of (8). We utilize the augmented Lagrangian method to solve constrained (9) by iteratively solving (11) and (12)

$$\begin{aligned}&(\boldsymbol{w}_{k+1},\boldsymbol{u}_{k+1},\boldsymbol{x}_{k+1}) \\ &\quad = \operatorname*{argmin}_{\boldsymbol{w},\boldsymbol{u},\boldsymbol{x}} \mathcal{L}_{\mathcal{A}}(\boldsymbol{w},\boldsymbol{u},\boldsymbol{x}),\end{aligned} \tag{11}$$

$$\begin{cases}\boldsymbol{\nu}_{k+1} = \boldsymbol{\nu}_k - \beta(\mathcal{D}\boldsymbol{u}_{k+1} - \boldsymbol{w}_{k+1}) \\ \boldsymbol{\gamma}_{k+1} = \boldsymbol{\gamma}_k - \theta(\boldsymbol{u}_{k+1} - \boldsymbol{x}_{k+1}) \\ \boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k - \mu(\boldsymbol{A}\boldsymbol{u}_{k+1} - \boldsymbol{b}).\end{cases} \tag{12}$$

Here, the subscript k denotes the iteration number, and $\boldsymbol{\nu}$, $\boldsymbol{\gamma}$, $\boldsymbol{\lambda}$ are the Lagrangian multipliers associated with the constraints $\mathcal{D}\boldsymbol{u} = \boldsymbol{w}$, $\boldsymbol{u} = \boldsymbol{x}$, $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{b}$, respectively.

In addition, (11) is still hard to solve efficiently in a direct way due to its nondifferentiability. Here, an alternating direction technique [32] is employed, which alternatively minimizes one variable while fixing the other variables, to split (11) into the following three subproblems. In the following, we argue that the every separated subproblem admits an efficient solution. For simplicity, the subscript k is omitted without confusion.

A. $\boldsymbol{w}$ Subproblem

Given $\boldsymbol{u}$, $\boldsymbol{x}$, after simplifications, the optimization problem associated with $\boldsymbol{w}$ can be expressed as

$$\begin{aligned}\min_{\boldsymbol{w}} \mathcal{Q}_1(\boldsymbol{w}) &= \min_{\boldsymbol{w}} \left\{\|\boldsymbol{w}\|_1 - \boldsymbol{\nu}^T(\mathcal{D}\boldsymbol{u} - \boldsymbol{w}) + \frac{\beta}{2}\|\mathcal{D}\boldsymbol{u} - \boldsymbol{w}\|_2^2\right\} \\ &= \min_{\boldsymbol{w}} \frac{1}{2}\left\|\left(\mathcal{D}\boldsymbol{u} - \frac{\boldsymbol{\nu}}{\beta}\right) - \boldsymbol{w}\right\|_2^2 + \frac{1}{\beta}\|\boldsymbol{w}\|_1.\end{aligned} \tag{13}$$

According to Lemma 1, the closed form of (13) is written as

$$\tilde{\boldsymbol{w}} = \max\left\{\left|\mathcal{D}\boldsymbol{u} - \frac{\boldsymbol{\nu}}{\beta}\right| - \frac{1}{\beta}, 0\right\} \cdot \operatorname{sgn}\left(\mathcal{D}\boldsymbol{u} - \frac{\boldsymbol{\nu}}{\beta}\right). \tag{14}$$

B. $\boldsymbol{u}$ Subproblem

With the aid of $\boldsymbol{w}$, $\boldsymbol{x}$, the $\boldsymbol{u}$ subproblem is equivalent to

$$\min_{\boldsymbol{u}} \mathcal{Q}_2(\boldsymbol{u}) = \min_{\boldsymbol{u}} \left\{\begin{array}{l}-\boldsymbol{\nu}^T(\mathcal{D}\boldsymbol{u} - \boldsymbol{w}) + \frac{\beta}{2}\|\mathcal{D}\boldsymbol{u} - \boldsymbol{w}\|_2^2 \\ -\boldsymbol{\gamma}^T(\boldsymbol{u} - \boldsymbol{x}) + \frac{\theta}{2}\|\boldsymbol{u} - \boldsymbol{x}\|_2^2 \\ -\boldsymbol{\lambda}^T(\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b}) + \frac{\mu}{2}\|\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b}\|_2^2\end{array}\right\}. \tag{15}$$

Clearly, $\mathcal{Q}_2(\boldsymbol{u})$ is a quadratic function and its gradient can be expressed as

$$\boldsymbol{d}(\mathcal{Q}_2(\boldsymbol{u})) = \mathcal{D}^T(\beta\mathcal{D}\boldsymbol{u} - \boldsymbol{\nu}) - \boldsymbol{\gamma} + \theta(\boldsymbol{u} - \boldsymbol{x}) - \beta\mathcal{D}^T\boldsymbol{w} + \boldsymbol{A}^T(\mu(\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b}) - \boldsymbol{\lambda}). \quad (16)$$

Setting $\boldsymbol{d}(\mathcal{Q}_2(\boldsymbol{u})) = 0$ gives us the exact minimizer of (15), that is

$$\boldsymbol{u} = \left(\beta\mathcal{D}^T\mathcal{D} + \theta\boldsymbol{I} + \mu\boldsymbol{A}^T\boldsymbol{A}\right)^+ \left(\mathcal{D}^T\boldsymbol{\nu} + \boldsymbol{\gamma} + \theta\boldsymbol{x} + \boldsymbol{A}^T\boldsymbol{\lambda} + \mu\boldsymbol{A}^T\boldsymbol{b} + \beta\mathcal{D}^T\boldsymbol{w}\right) \quad (17)$$

where $\mathbf{M}^+$ stands for the Moore–Penrose pseudoinverse of matrix $\mathbf{M}$. Theoretically, it is ideal to accept the exact minimizer as the solution of the $\boldsymbol{u}$ subproblem. However, computing the inverse or pseudoinverse at each iteration is too costly to implement numerically. Therefore, an iterative method is highly desirable. Here, the steepest descent method with the optimal step is used to solve (15) iteratively by applying

$$\tilde{\boldsymbol{u}} = \boldsymbol{u} - \eta\boldsymbol{d} \quad (18)$$

where $\boldsymbol{d}$ is the gradient direction of the objective function $\mathcal{Q}_2(\boldsymbol{u})$, $\eta = \mathbf{abs}(\boldsymbol{d}^T\boldsymbol{d}/\boldsymbol{d}^T\boldsymbol{G}\boldsymbol{d})$ represents the optimal step, $\boldsymbol{G} = (\beta\mathcal{D}^T\mathcal{D} + \theta\boldsymbol{I} + \mu\boldsymbol{A}^T\boldsymbol{A})$, and $\boldsymbol{I}$ is the identity matrix. Therefore, solving $\boldsymbol{u}$ subproblem requires computing (18) from iteration to iteration.

C. $\boldsymbol{x}$ Subproblem

Given $\boldsymbol{w}$, $\boldsymbol{u}$, similarly, the $\boldsymbol{x}$ subproblem becomes

$$\min_{\boldsymbol{x}} \mathcal{Q}_3(\boldsymbol{x}) = \min_{\boldsymbol{x}} \left\{ \alpha\|\boldsymbol{\Theta}_{\boldsymbol{x}}\|_0 - \boldsymbol{\gamma}^T(\boldsymbol{u} - \boldsymbol{x}) + \frac{\theta}{2}\|\boldsymbol{u} - \boldsymbol{x}\|_2^2 \right\}. \quad (19)$$

After making a straightforward complete-the-squares procedure and omitting a constant independent of $\boldsymbol{x}$, (19) can be written as

$$\min_{\boldsymbol{x}} \mathcal{Q}_3(\boldsymbol{x}) = \min_{\boldsymbol{x}} \left\{ \frac{\theta}{2}\|\boldsymbol{x} - r\|_2^2 + \alpha\|\boldsymbol{\Theta}_{\boldsymbol{x}}\|_0 \right\} = \min_{\boldsymbol{x}} \left\{ \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{r}\|_2^2 + \frac{\alpha}{\theta}\sum_{k=1}^{n} \left\|\boldsymbol{T}^{\boldsymbol{3D}}(\boldsymbol{Z}_{\boldsymbol{x}_k})\right\|_0 \right\} \quad (20)$$

where $\boldsymbol{r} = (\boldsymbol{u} - (\boldsymbol{\gamma}/\theta))$.

Note that it is difficult to solve (20) directly due to the complicated definition of $\boldsymbol{\Theta}_{\boldsymbol{x}}$. To enable solving (20) tractable, in this paper, a general assumption is made, with which even a closed form of (20) can be achieved. Concretely, we regard $\boldsymbol{r}$ as some type of the noisy observation of $\boldsymbol{x}$, denote the error vector by $\boldsymbol{e} = \boldsymbol{x} - \boldsymbol{r}$, and then make an assumption that each element of $\boldsymbol{e}$ follows an independent zero-mean distribution with the same variance σ^2. It is worth emphasizing that the above assumption does not need to be Gaussian process, which is more general and reasonable. Based on the assumption, since $\boldsymbol{e}, \boldsymbol{x}, \boldsymbol{r} \in \mathbb{R}^N$, by invoking the *Law of Large Numbers* in probability theory, we have the following equation with very large probability (limited to 1)

$$\frac{\|\boldsymbol{e} - \boldsymbol{0}\|_2^2}{N} = \frac{\|\boldsymbol{e}\|_2^2}{N} = \frac{\|\boldsymbol{x} - \boldsymbol{r}\|_2^2}{N} = \sigma^2. \quad (21)$$

TABLE II
CS RECOVERY VIA COLLABORATIVE SPARSITY

Input: The observed measurement $\boldsymbol{b}$, the measurement matrix $\boldsymbol{A}$ and β, μ, θ, τ.
Initialization: $\boldsymbol{u}_0 = \boldsymbol{A}^T\boldsymbol{b}$, , $\mathbf{v}_0 = \boldsymbol{\gamma}_0 = \boldsymbol{\lambda}_0 = \mathbf{0}$; $\boldsymbol{w}_0 = \boldsymbol{x}_0 = \mathbf{0}$
while Outer stopping criteria unsatisfied ***do***
while Inner stopping criteria unsatisfied ***do***
Solve $\boldsymbol{w}$ sub-problem by computing Eq. (14);
Solve $\boldsymbol{u}$ sub-problem by computing Eq. (18);
Solve $\boldsymbol{x}$ sub-problem by computing Eq. (24);
end while
Update multipliers $\mathbf{v}, \boldsymbol{\gamma}, \boldsymbol{\lambda}$ by computing Eq. (12);
end while
Output: Final restored image $\tilde{\boldsymbol{u}}$.

At the same time, since the orthogonal transform $\boldsymbol{T}^{\boldsymbol{3D}}$ for every group has the property of energy conservation and $\boldsymbol{\Theta}_{\boldsymbol{x}}$, $\boldsymbol{\Theta}_{\boldsymbol{r}} \in \mathbb{R}^K$, according to the *Law of Large Numbers*, there also exists the following equation with very large probability:

$$\frac{\|\boldsymbol{\Theta}_{\boldsymbol{x}} - \boldsymbol{\Theta}_{\boldsymbol{r}}\|_2^2}{K} = \frac{\sum_{k=1}^{n} \left\|\boldsymbol{T}^{\boldsymbol{3D}}(\boldsymbol{Z}_{\boldsymbol{x}_k}) - \boldsymbol{T}^{\boldsymbol{3D}}(\boldsymbol{Z}_{\boldsymbol{r}_k})\right\|_2^2}{K} = \sigma^2. \quad (22)$$

Incorporating (22) into (21) leads to

$$\min_{\boldsymbol{x}} \mathcal{Q}_3(\boldsymbol{x}) = \min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{\Theta}_{\boldsymbol{x}} - \boldsymbol{\Theta}_{\boldsymbol{r}}\|_2^2 + \frac{K\alpha}{N\theta}\|\boldsymbol{\Theta}_{\boldsymbol{x}}\|_0. \quad (23)$$

Owing to Lemma 2, the closed form of (23) is written as

$$\tilde{\boldsymbol{\Theta}}_{\boldsymbol{x}} = \mathbf{hard}\left(\boldsymbol{\Theta}_{\boldsymbol{r}}, \sqrt{2\tau}\right) = \boldsymbol{\Theta}_{\boldsymbol{r}} \cdot \mathbf{1}\left(\mathbf{abs}(\boldsymbol{\Theta}_{\boldsymbol{r}}) - \sqrt{2\tau}\right)$$

where $\tau = K\alpha/N\theta$. Thus, the efficient solution for the $\boldsymbol{x}$ (19) is

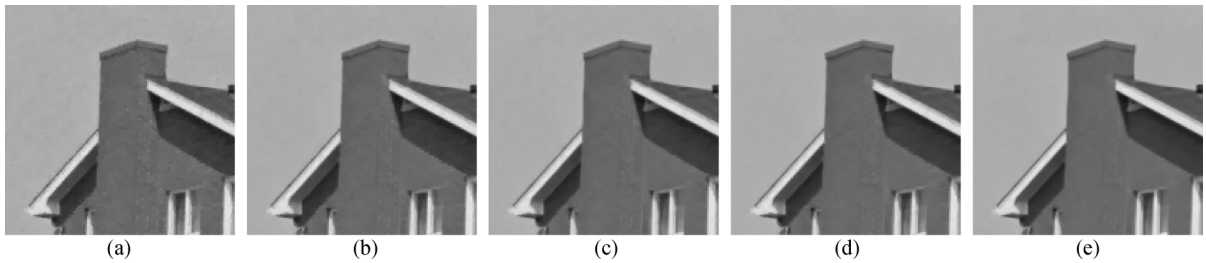
$$\tilde{\boldsymbol{x}} = \boldsymbol{\Omega}_{\boldsymbol{N3D}}\tilde{\boldsymbol{\Theta}}_{\boldsymbol{x}} = \boldsymbol{\Omega}_{\boldsymbol{N3D}}\left(\mathbf{hard}\left(\boldsymbol{\Theta}_{\boldsymbol{r}}, \sqrt{2\tau}\right)\right). \quad (24)$$

D. Summary of RCoS

So far, all issues in the process of handing the subproblems have been solved. In fact, we achieve the efficient solution for each separated subproblem, which will enable the whole algorithm more efficient and effective. In light of all derivations above, the complete description of RCoS for CS-acquired images is composed of two loops: inner loop and outer loop, as stated clearly in Table II.

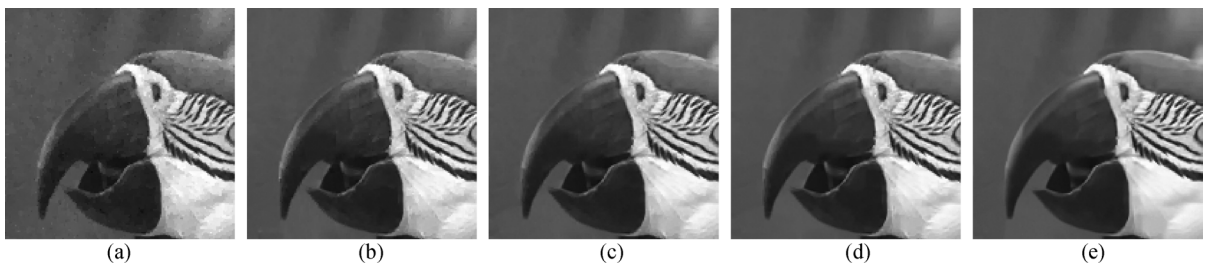
V. EXPERIMENTAL RESULTS

In this section, extensive experimental results are presented to evaluate the performance of the proposed RCoS. For thoroughness and fairness of our comparative study, we exploit a broad class of natural images, including six conventional images (256×256) and two biomedical images (112×112). In our experiments, the CS measurements are obtained by applying a random projection matrix to the original image signal, where the random projection matrix is generated by Matlab command



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

Fig. 3. CS recovered results by RCoS with various choices of $\beta = 32$ and $\tau = 8$ (30% measurements). (a) Recovered by TV [16], PSNR $= 36.17$ dB. (b) $\beta = 32$ and $\tau = 8$, PSNR $= 37.23$ dB. (c) $\beta = 32$ and $\tau = 12$, PSNR $= 37.31$ dB. (d) $\beta = 64$ and $\tau = 8$, PSNR $= 37.48$ dB. (e) $\beta = 64$ and $\tau = 12$, PSNR $= 37.29$ dB.



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

Fig. 4. CS recovered results by RCoS with various choices of $\beta = 32$ and $\tau = 8$ (30% measurements). (a) Recovered by TV [16], PNSR $= 30.83$ dB. (b) $\beta = 32$ and $\tau = 8$, PNSR $= 37.78$ dB. (c) $\beta = 32$ and $\tau = 12$, PNSR $= 32.99$ dB. (d) $\beta = 64$ and $\tau = 8$, PNSR $= 32.84$ dB. (e) $\beta = 64$ and $\tau = 12$, PNSR $= 32.45$ dB.

$\mathbf{rand}(\mathbf{M}, \mathbf{N})$. To save memory required to store the projection matrix, we first split an image with size 256×256 into four 128×128 subimages, conduct the CS recovery for each subimage, respectively, and merge them into a whole recovered image at last.

A. Parameter Setting

All the experiments are performed in Matlab 7.12.0 on a Dell OPTIPLEX computer with Intel Core 2 Duo CPU E8400 processor (3.00 GHz), 3.25 G memory, and Windows XP operating system. The 3-D transform denoted by $\boldsymbol{T^{3D}}$ is 3-D wavelet transform, composed of 2-D bior1.5 and 1-D Haar. The way for the stopping criteria in Table II is the relative change of $\boldsymbol{u}$ is sufficiently small, i.e., $\|\boldsymbol{u}_{k+1} - \boldsymbol{u}_k\|_2^2 \leq \varepsilon$ and ε is set to be 10^{-3}.

In our implementation, all the parameters of RCoS including B_s, L, c, μ, θ, β, τ are set empirically, where B_s, L, c, μ, θ are set fixed for all test images and β, τ can be changed according to different images. Concretely, the size of each block, i.e., B_s is set to be 8×8, the size of training window for searching matched blocks, i.e., $L \times L$ is set to be 41×41, and the number of best matched blocks, i.e., c is set to be 10. We empirically set $\mu = 128$ and $\theta = 2$. From our experiments, a reasonable β and τ for a natural image lies in the range [32, 64] and [8, 12], respectively. Figs. 3 and 4 show the recovery results by using different values β and τ.

From Figs. 3 and 4, it can be seen that the reconstruction results obtained by the four combinations of β and τ [see Fig. 3 and Fig. 4(b)–(e)] are very close in visual perception, and are all better than the results by TV method [Fig. 3 and Fig. 4(a)]. In practice, for any natural image, the default setting of β and τ can be $\beta = 64$ and $\tau = 8$. Here, in this paper, we set $\beta = 32$ or 64, and $\tau = 8$ or 12, and chose the optimal combination for the best image quality to demonstrate the performance of the proposed algorithm. It is fair since all the results of other comparative methods (TV [16], TSW [18], TSDCT [19]) are all generated by the original authors' codes with the corresponding parameters manually optimized. Further study on determining the regularization parameters of the proposed RCoS automatically is our future work.

B. Experimental Comparisons

RCoS is compared with three representative image CS recovery methods in literatures, i.e., tree-structured wavelet (TSW) method [18], tree-structured DCT (TSDCT) method [19] and total variation (TV) method [16], which deal with the image signal in the wavelet domain, the DCT domain and the gradient domain, respectively. It is worth emphasizing that total variation (TV) method is known as one of the state-of-the-art algorithms for image CS recovery. PSNR, a classical quality metric, is used to evaluate the results obtained by different methods. The RCoS software and more experimental results can be found online.[1]

The PSNR results reconstructed by various methods for all test images are provided in Table III. TSW [18] obtains the lowest PSNR among the four comparative algorithms. The proposed RCoS considerably outperforms the other methods in all the cases, with PSNR improvements of up to 5.3, 12.6, and 12.8 dB, compared with TV, TSDCT, and TSW, respectively.

Furthermore, we compare the average performance of these four methods. For all test images, each average PSNR result with regarding to every ratio of measurement generated by each comparative algorithm is plot in Fig. 5. It is clear to see that

[1]Available online: http://idm.pku.edu.cn/staff/zhangjian/RCoS/.

TABLE III
PSNR RESULTS OF FOUR CS RECOVERY ALGORITHMS (UNIT: DB)

Image	***Barbara*** (256×256)				***Leaves*** (256×256)				***Vessels*** (112×112)				***Parrots*** (256×256)			
ratio (%)	15	20	25	30	15	20	25	30	15	20	25	30	15	20	25	30
TSW[18]	21.06	21.51	21.82	22.35	16.94	17.71	18.56	19.33	17.71	19.16	20.16	21.07	23.46	24.61	25.87	27.52
TSDCT[19]	21.21	22.87	23.65	25.44	17.66	19.97	20.83	22.81	19.68	22.64	23.81	26.62	24.97	27.92	29.09	31.55
TV[16]	23.52	24.42	25.18	26.31	21.28	23.08	24.64	26.61	22.54	24.56	26.12	27.94	27.58	29.44	31.17	32.21
RCoS	25.56	27.67	29.60	31.44	25.67	28.18	29.63	31.81	26.02	28.81	30.69	33.24	30.48	32.33	34.03	34.52
Image	***Chest*** (112×112)				***Lena*** (256×256)				***House*** (256×256)				***Cameraman*** (256×256)			
ratio (%)	15	20	25	30	15	20	25	30	15	20	25	30	15	20	25	30
TSW[18]	17.26	19.56	20.97	26.13	23.29	24.27	25.16	25.97	26.57	27.95	29.81	31.05	21.35	23.08	24.86	25.70
TSDCT[19]	18.68	20.77	21.17	23.60	23.86	26.43	26.85	28.75	28.82	31.40	31.83	33.47	22.77	24.79	25.17	27.35
TV[16]	25.08	28.76	31.91	34.45	27.45	28.69	30.00	31.14	31.51	33.05	34.36	34.96	25.73	27.74	28.98	30.26
RCoS	28.25	31.35	33.77	35.99	29.16	30.86	32.37	33.58	34.24	35.26	36.12	36.30	28.61	29.74	30.77	31.70

RCoS gains as much as about 3 dB over the second best recovery method (i.e., TV), 6 dB over TSDCT and 8 dB over TSW. Additionally, as shown in Fig. 5, the reconstruction quality achieved by RCoS in the case of ratio $= 15\%$ is still higher than the one by TV in the case of ratio $= 20\%$ and the one by TSDCT in the case of ratio $= 30\%$.

Some visual results of the recovered images for the four algorithms are presented in Figs. 6–10, which verify the superiority of RCoS in preserving the image structures and fine details, showing much better visual results than the other competing methods. Moreover, we use a new image quality assessment (IQA) model FSIM to evaluate the visual quality, which is proposed recently and achieves much higher consistency with the subjective evaluations than state-of-the-art IQA metrics [34]. The higher FSIM value means the better visual quality. It could be seen that proposed RCoS achieves the highest FSIM scores in all the figures, which again demonstrates that RCoS can achieve better performance on the image visual quality.

The high performance of RCoS is attributed to the proposed adaptive hybrid space-transform domain, which offers a powerful mechanism of characterizing the structured sparsities of natural image signals.

To investigate the robustness of the proposed algorithm, we further consider a CS recovery problem with Gaussian noise. That means the measurement observation $\boldsymbol{b}$ in (1) becomes

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{n}$$

where $\boldsymbol{n}$ has mean zero and standard deviation σ.

Table IV shows the PSNR results of the four test images in 20% and 30% measurements with Gaussian noise under different standard deviations from 50 to 200. It can be observed that TV and RCoS are both robust to the observation Gaussian noise, while the recovered results by RCoS significantly outperforms the ones by TV.

C. Convergence Study

Because the objective function (1), comprising mixed ℓ_0 and ℓ_1 norms, is nonconvex, it is difficult to give its theoretical proof for convergence. In this subsection, we only provide empirical evidence to illustrate the convergence of the proposed RCoS and show some important results. Fig. 11 plots the evolutions of PSNR versus iteration numbers for Image *Vessels* with various ratios of measurements.

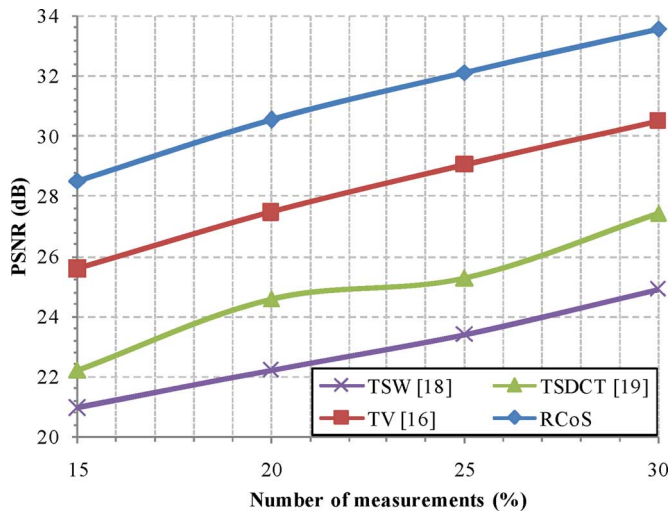
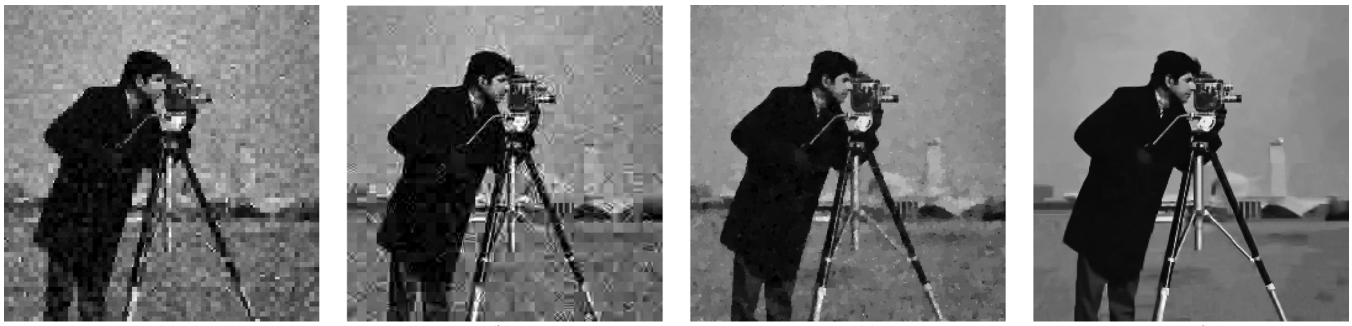


Fig. 5. Average PSNR result achieved by each algorithm for all test images versus different numbers of measurements.

It is observed that with the increase of iteration number, the PSNRs of the reconstructed image increase greatly at first and then tend to be stable. Another observation is that the rising curves in Fig. 11 are not smooth and have many sharp points about every one hundred iterations. There are probably two reasons, which lead to the above phenomenon. One is the inherent nonconvex property the problem possesses. The other manifests that the proposed RCoS has the ability to avoid converging local optimal solution, trying to achieve the global optimal solution.

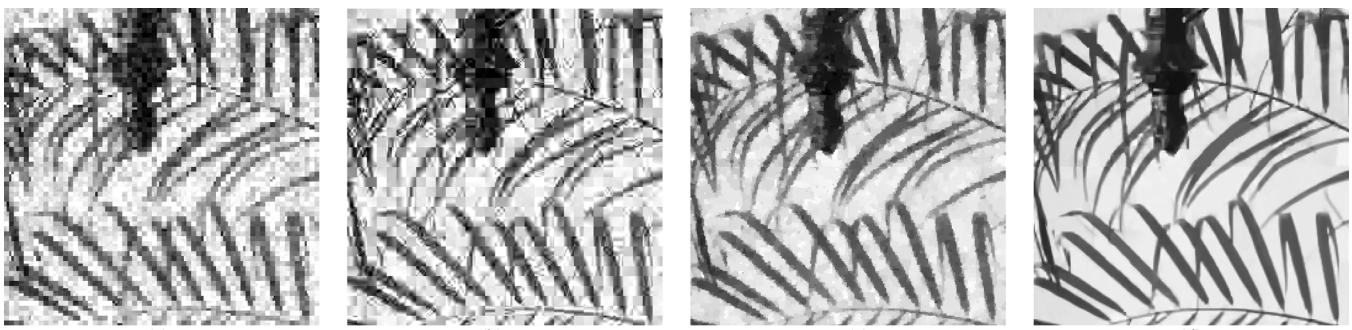
D. Computational Time

Finally, we provide computational time comparison for all the test images between TV [16] and the proposed RCoS in Table V. On average, the time computational complexity of RCoS is rough 2.2 times that of TV. Comparing the $\boldsymbol{w}$, $\boldsymbol{u}$, $\boldsymbol{x}$ problems described in Table II, it is obvious to conclude that



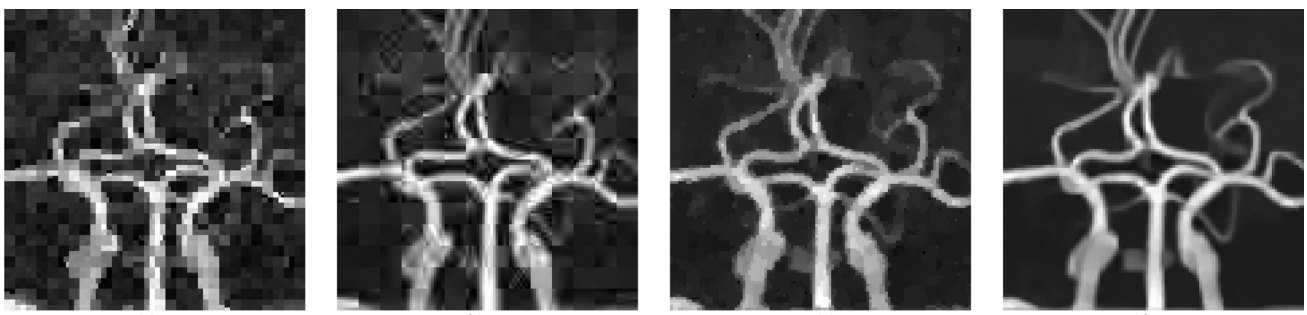
(a) (b) (c) (d)

Fig. 6. CS recovered image *cameraman* (15% measurements). (a) TSW (21.35 dB, FSIM $= 0.7042$). (b) TSDCT (22.77 dB, FSIM $= 0.7487$). (c) TV (25.73 dB, FSIM $= 0.8247$). (d) RCoS (28.61 dB, FSIM $= 0.8618$).



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

Fig. 7. CS recovered image *leaves* (15% measurements). (a) TSW (16.94 dB, FSIM $= 0.6280$). (b) TSDCT (17.66 dB, FSIM $= 0.6723$). (c) TV (21.28 dB, FSIM $= 0.8038$). (d) RCoS (25.67 dB, FSIM $= 0.9002$).



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

Fig. 8. CS recovered image *vessels* (20% measurements). (a) TSW (19.16 dB, FSIM $= 0.7314$). (b) TSDCT (22.64 dB, FSIM $= 0.8206$). (c) TV (24.56 dB, FSIM $= 0.8700$). (d) RCoS (28.81 dB, FSIM $= 0.9321$).

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

Fig. 9. CS recovered image *parrots* (20% measurements). (a) TSW (24.61 dB, FSIM $= 0.7939$). (b) TSDCT (27.92 dB, FSIM $= 0.8719$). (c) TV (29.44 dB, FSIM $= 0.8927$). (d) RCoS (32.33 dB, FSIM $= 0.9307$).

the main complexity of RCoS comes from the $\boldsymbol{x}$ subproblem, which needs similar block searching and the operations of 3-D and inverse 3-D transforms.

To speed up RCoS, on one hand, we can exploit the results of TV instead of zeros as initialization, which could decrease the number of iteration enormously. On the other hand, ongoing

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

Fig. 10. CS recovered image *Barbara* (25% measurements). (a) TSW (21.82 dB, FSIM = 0.7481). (b) TSDCT (23.65 dB, FSIM = 0.8231). (c) TV (25.18 dB, FSIM = 0.8403). (d) RCoS (29.60 dB, FSIM = 0.9123).

TABLE IV
PSNR RESULTS OF TWO CS RECOVERY ALGORITHMS WITH GAUSSIAN NOISE (UNIT: DB)

Image	***Barbara*** (256×256)				***Leaves*** (256×256)				***Vessels*** (112×112)				***Parrots*** (256×256)			
σ	50	100	150	200	50	100	150	200	50	100	150	200	50	100	150	200
ratio (%)	20				20				20				20			
TV [16]	24.40	24.34	24.25	24.12	23.05	22.99	22.90	22.77	24.47	24.26	23.99	23.66	29.39	29.22	28.94	28.57
RCoS	26.89	26.63	26.40	26.27	28.00	27.78	27.51	27.25	28.43	28.12	27.88	27.13	32.01	31.73	31.57	31.09
ratio (%)	30				30				30				30			
TV [16]	26.28	26.17	25.99	25.74	26.56	26.39	26.14	25.81	27.72	27.23	26.55	25.84	32.07	31.67	31.06	30.34
RCoS	30.89	30.46	29.88	29.66	31.41	31.23	30.75	30.13	32.48	31.65	30.81	29.55	34.31	33.86	33.16	32.51

TABLE V
COMPUTATIONAL TIME COMPARISON OF TWO CS RECOVERY ALGORITHMS (UNIT: S)

Image	***Barbara*** (256×256)				***Leaves*** (256×256)				***Vessels*** (112×112)				***Parrots*** (256×256)			
ratio (%)	15	20	25	30	15	20	25	30	15	20	25	30	15	20	25	30
TV [16]	415	498	545	604	435	430	585	655	33	39	43	46	432	514	555	608
RCoS	1048	1284	1507	1585	984	1035	977	994	86	100	115	128	892	960	1060	1164
Image	***Chest*** (112×112)				***Lena*** (256×256)				***House*** (256×256)				***Cameraman*** (256×256)			
ratio (%)	15	20	25	30	15	20	25	30	15	20	25	30	15	20	25	30
TV [16]	28	32	33.8	36	424	497	536	597	418	489	518	563	490	575	625	680
RCoS	69	79	90.7	101	1010	1255	1280	1520	683	795	1013	1187	975	1136	1259	1418

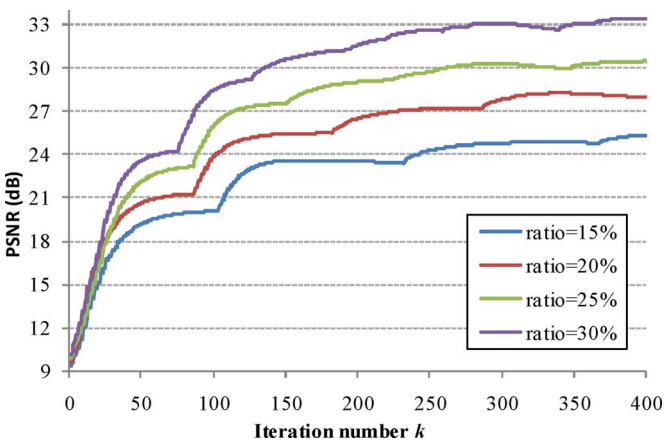


Fig. 11. Progression of the PSNR (dB) of the reconstructed Image *vessels* with respect to the iteration count k in the cases of different ratios of measurements.

work addresses the parallelization, leveraging GPU hardware to speed up RCoS.

VI. CONCLUSION

In this paper, a novel sparsity measure, called collaborative sparsity measure is introduced, and a new strategy for CS RCoS is proposed, which efficiently characterizes the intrinsic sparsities of natural images in an adaptive hybrid space-transform domain. Extensive experiments on a wide range of CS-acquired images manifest that RCoS is able to increase recovery quality by a large margin compared with the current methods or require many fewer measurements for a desired reconstruction quality. Our work offers a fresh and successful instance to corroborate the CS theory applied for real signals (i.e., natural images).

ACKNOWLEDGMENT

The authors would like to thank the authors of [16], [18], and [19] for kindly providing their code. The authors would would also like to thank the anonymous reviewers for their constructive suggestions which helped us improving our manuscript.

REFERENCES

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

[2] 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.

[3] E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[4] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[5] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE J. Sel. Topic Signal Process.*, vol. 1, no. 4, pp. 568–597, Dec. 2007.

[6] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–1666, Dec. 2008.

[7] I. Daubechies, M. D. Friese, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.

[8] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l1 minimization," *J. Fourier Anal. Applicat.*, vol. 14, no. 5–6, pp. 877–905, Dec. 2008.

[9] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[10] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace prior," *IEEE Trans. Image Process.*, vol. 19, no. 1, pp. 53–63, Jan. 2010.

[11] W. Guo and W. Yin, "Compressed sensing via iterative support detection," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 462–491, Aug. 2010.

[12] D. Baron, S. Sarvotham, and R. G. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 269–280, Jan. 2010.

[13] X. Wu, X. Zhang, and J. Wang, "Model-guided adaptive recovery of compressive sensing," in *Proc. IEEE Data Compress. Conf.*, 2009, pp. 123–132.

[14] *YALL1 Toolbox*, [Online]. Available: http://yall1.blogs.rice.edu/

[15] ℓ_1*-Magic Toolbox*, [Online]. Available: http://users.ece.gatech.edu/~justin/l1magic/

[16] C. Li, W. Yin, and Y. Zhang, TVAL3: TV minimization by augmented Lagrangian and alternating direction agorithm 2009 [Online]. Available: http://www.caam.rice.edu/ ~optimization/L1/TVAL3/

[17] Y. Kim, M. S. Nadar, and A. Bilgin, "Compressed sensing using a Gaussian scale mixtures model in wavelet domain," in *Proc. IEEE Int. Conf. Image Process.*, 2010, pp. 3365–3368.

[18] L. He and L. Carin, "Exploiting structure in wavelet-based Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3488–3497, Sep. 2009.

[19] L. He, H. Chen, and L. Carin, "Tree-structured compressive sensing with variational Bayesian analysis," *IEEE Signal Process. Lett.*, vol. 17, no. 3, pp. 233–236, Mar. 2010.

[20] S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," in *Proc. IEEE Int. Conf. Image Process.*, 2009, pp. 3021–3024.

[21] E. Candes and T. Tao, "Near-optimal signal recovery from random projections and universal encoding strategies?," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5245, Dec. 2006.

[22] M. Duarte, M. Davenport, D. Takhar, J. Laska, and T. Sun, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[23] J. Nocedal and S. J. Wright*, Numerical Optimization*, 2nd ed. New York: Springer, 2006.

[24] M. Afonso, J. Bioucas-Dias, and M. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.

[25] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, vol. 22, pp. 1–9.

[26] M. K. Varanasi and B. Aazhang, "Parametric generalized Gaussian density estimation," *J. Acoust. Soc. Am.*, vol. 86, no. 4, pp. 1404–1415, 1989.

[27] S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vis.*, vol. 82, no. 2, pp. 205–229, 2009.

[28] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 60–65.

[29] 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.

[30] 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.

[31] 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, 2009.

[32] J. F. Cai, S. Osher, and Z. W. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Model. Simul.*, pp. 5057–5071, 2009.

[33] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, and W. Gao, "Compressed sensing recovery via collaborative sparsity," in *Proc. IEEE Data Compress. Conf.*, Apr. 2012, pp. 287–296.

[34] 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.

[35] G. Peyre, S. Bougleux, and L. Cohen, "Non-local regularization of inverse problems," in *Proc. Eur. Conf. Comput. Vis.*, Marseille, France, Oct. 2008.

[36] Y. H. Xiao and H. N. Song, "An inexact alternating directions algorithm for constrained total variation regularized compressive sensing problems," *J. Math. Imag. Vis.*, vol. 44, no. 2, pp. 114–127, 2012.

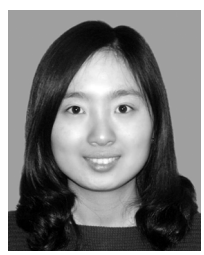
Jian Zhang (S'12) received the B.S. and M.S. degrees, in 2007 and 2009, respectively, from Department of Mathematics and School of Computer Science and Technology, Harbin Institute of Technology (HIT), Harbin, China, where he is currently working toward the Ph.D. degree in School of Computer Science and Technology.

In 2008, he joined the Joint Research and Development Lab (JDL), Chinese Academy of Sciences, Beijing, as a Research Assistant. Since September 2009, he has been with the National Engineering Lab for Video Technology, Peking University, Beijing, China, as a Research Assistant. His research interests include image/video processing, compressive sensing, sparse representation and dictionary learning.

Mr. Zhang was the recipient of the Best Paper Award at IEEE Visual Communication and Image Processing (VCIP) 2011.

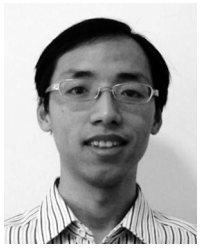
Debin Zhao (M'11) received the B.S., M.S., and Ph.D. degrees in computer science from Harbin Institute of Technology (HIT), Harbin, China, in 1985, 1988, and 1998, respectively.

He is now a Professor in the Department of Computer Science, Harbin Institute of Technology. He has published over 200 technical articles in refereed journals and conference proceedings in the areas of image and video coding, video processing, video streaming and transmission, and pattern recognition.



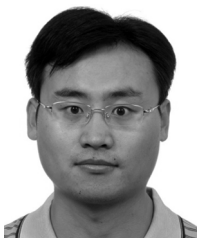
Chen Zhao received the B.S. degree from Si Chuan University, Cheng Du, China, in 2010. She is now working toward the Ph.D. degree at School of Electrical Engineering and Computer Science, Peking University, Beijing, China.

Her research interests include image/video processing and video coding.



Ruiqin Xiong (M'08) received the B.S. degree from University of Science and Technology of China, Hefei, China, in 2001, and the Ph.D. degree from Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2007.

He was a Research Intern at Microsoft Research Asia from August 2002 to July 2007, and a Senior Research Associate at University of New South Wales, Australia, from September 2007 to August 2009. Then he joined Peking University, Beijing, China, where he is now a Research Professor. His research interests include image and video processing, compression, multimedia communication, channel coding and distributed coding, etc.



Siwei Ma (M'05) received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, in 2005.

From 2005 to 2007, he was a post-doctorate with the University of Southern California, Los Angeles. Then he joined the Institute of Digital Media, School of Electrical Engineering and Computer Science, Peking University, Beijing, China, where he is currently an Associate Professor. He has published over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.

Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991.

He is a Professor of computer science at Peking University, Beijing, China. Before joining Peking University, he was a Professor of computer science at the Harbin Institute of Technology from 1991 to 1995, and a Professor at the Institute of Computing Technology of Chinese Academy of Sciences. He has published extensively including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. He served or serves on the editorial board for EURASIP *Journal of Image Communications* and the *Journal of Visual Communication and Image Representation*.

Dr. Gao served or serves on the editorial board for several journals, such as IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE TRANSACTIONS ON MULTIMEDIA, and IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.