Hindawi
Mathematical Problems in Engineering
Volume 2018, Article ID 7171352, 15 pages
https://doi.org/10.1155/2018/7171352

Research Article

Content-Aware Compressive Sensing Recovery Using Laplacian Scale Mixture Priors and Side Information

Zhonghua Xie[1], **Lihong Ma**[1], **and Lingjun Liu**[1,2]

[1]*School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China*
[2]*School of Electronic and Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China*

Correspondence should be addressed to Lihong Ma; eelhma@scut.edu.cn

Received 10 August 2017; Revised 3 November 2017; Accepted 20 November 2017; Published 29 January 2018

Academic Editor: Raffaele Solimene

Nonlocal methods have shown great potential in many image restoration tasks including compressive sensing (CS) reconstruction through use of image self-similarity prior. However, they are still limited in recovering fine-scale details and sharp features, when rich repetitive patterns cannot be guaranteed; moreover the CS measurements are corrupted. In this paper, we propose a novel CS recovery algorithm that combines nonlocal sparsity with local and global prior, which soften and complement the self-similarity assumption for irregular structures. First, a Laplacian scale mixture (LSM) prior is utilized to model dependencies among similar patches. For achieving group sparsity, each singular value of similar packed patches is modeled as a Laplacian distribution with a variable scale parameter. Second, a global prior and a compensation-based sparsity prior of local patch are designed in order to maintain differences between packed patches. The former refers to a prediction which integrates the information at the independent processing stage and is used as side information, while the latter enforces a small (i.e., sparse) prediction error and is also modeled with the LSM model so as to obtain local sparsity. Afterward, we derive an efficient algorithm based on the expectation-maximization (EM) and approximate message passing (AMP) frame for the maximum a posteriori (MAP) estimation of the sparse coefficients. Numerical experiments show that the proposed method outperforms many CS recovery algorithms.

1. Introduction

Compressive sensing (CS) [1, 2] allows us to reconstruct high-dimensional data with only a small number of random samples or measurements, if the original signal can be sparsely represented by some given appropriate basis. Owing to the fact that image prior knowledge plays a critical role in the performance of compressive sensing reconstruction, much efforts have been made to develop an effective regularization term or signal model to reflect the image prior knowledge. Standard CS methods exploit the sparsity of signal in some domains, such as DCT [3], wavelets [4, 5], total variation (TV) [6, 7], and learned dictionary [8, 9]. Unfortunately, these methods are less appropriate for many imaging applications. The reason for this failure is that natural images do not have an exactly sparse representation in any above basis. These models favor piecewise constant image structures and hence tend to smooth much the image details.

More recently, the concept of sparsity has evolved into various sophisticated forms, including group sparsity [10, 11], tree sparsity [12–14], and nonlocal sparsity [15–19], where higher-order dependency among sparse coefficients is exploited. Among them, nonlocal sparsity, which refers to the fact that a patch often has many nonlocal similar patches to it across the image, has been shown to be most beneficial to CS image recovery. In [15], a nonlocal total variation (NLTV) regularization model for CS image recovery is proposed by using the self-similarity property in gradient domain. In order to obtain an adaptive sparsity regularization term for CS image recovery process, a local piecewise autoregressive model is designed in [16]. In [17], similar patches are grouped to form a two-dimensional data matrix for characterizing the low-rank property, leading to a CS recovery method via nonlocal low-rank regularization (NLR-CS). In [18, 19], a probabilistic graphical model is established, which uses collaborative filtering [20] to promote sparsity of packed patches.

Despite the steady progress in nonlocal methods, they still tend to smooth the detailed image textures, degrading the image visual quality, for the reason that the lack of self-repetitive structures and corruption for data is unavoidable.

To deal with this issue, local and global priors are designed to soften and complement the nonlocal sparsity for irregular structures so as to preserve image details. More specifically, the nonlocal sparsity is only imposed on a set of patches with limited influence from neighboring pixels, while the global prior refers to a prediction used as a reference which integrates the outcomes at the independent processing stage and can maintain the entire consistency of image; moreover, a compensation-based constraint term of local patch is utilized to enforce a small (i.e., sparse) prediction error. Both local sparsity and nonlocal sparsity are represented by Laplacian scale mixture (LSM) [21, 22] models, which are adopted to force coefficients, that is, singular values of local patches and similar packed patches, to be sparse. Each coefficient is modeled as a Laplacian distribution with a variable scale parameter, resulting in weighted singular value minimization problems, where weights are adaptively assigned according to the signal-to-noise ratio. On the other hand, the reference image can be used as side information. Finally, we obtain a side information-aided LSM prior model for CS image reconstruction. To solve this model, the expectation-maximization (EM) [23] method is adopted, turning the CS recovery problem into a prior parameter estimation problem and a singular value minimization problem. In particular, owing to its promising performance and efficiency, we are motivated to apply the approximate message passing (AMP) algorithm [24, 25], which is an iterative algorithm that can be used in signal and image reconstruction by performing denoising at each iteration, to solve the latter. Experimental results on natural images show that our approach can achieve more accurate reconstruction than other competing approaches.

2. Background

2.1. CS Recovery Problem. The CS recovery problem aims to find the sparsest solution $x \in \mathbb{C}^n$ from the underdetermined linear system $y = Ax + w$, where $y \in \mathbb{C}^m$ are the measurements, $A \in \mathbb{C}^{m\times n}$, $m < n$ is the measurement matrix, and w denotes the additive noise. One can solve the following objective function:

$$\widehat{x} = \arg\min_x \frac{\|y - Ax\|_2^2}{2} + \lambda \mathfrak{R}(x). \tag{1}$$

The first term is the data fidelity term that represents the closeness of the solution to the measurements. The second term is a regularization term that represents a priori sparse information of the original signal. λ is a regularization parameter that balances the contribution of both terms. As mentioned in Introduction, CS recovery methods exploit the sparsity of signal in some domains, such as DCT [3], wavelets [4, 5], learned dictionary [8, 9], and total variation (TV) [6, 7], leading to various forms of $\mathfrak{R}(x) : \|\mathrm{DCT}(x)\|_p$, $\|\Psi(x)\|_p$, and $\|\mathrm{Dic}(x)\|_p$ where $p \in \{0, 1\}$ and $\|x\|_{\mathrm{TV}}$, respectively.

2.2. Nonlocal Sparsity. The abundance of self-repeating patterns in natural image (as shown in Figure 1) can be characterized by the nonlocal sparsity [16, 17]. As shown in Figure 2, for each local patch, we can find the first M most similar nonlocal patches to it. In practice, this can be done by Euclidean distance based block matching in a large enough local window. Let $R_i x$ (or x_i) denote an exemplar patch located at the ith position. Patches that are similar to $R_i x$ including $R_i x$ itself are found to form a low-rank matrix $C_i x = [R_{i_1} x, R_{i_2} x, \ldots, R_{i_N} x]$: $C_i x \in \mathbb{R}^{M\times N}$. Here, we suppose that $M \geq N$. An objective function that reflects the group sparsity of similar patches with a low-rank regularization term for CS recovery can be formulated as follows:

$$\widehat{x} = \arg\min_x \frac{\|y - Ax\|_2^2}{2} + \sum_{i=1}^{G} \|C_i x\|_*, \tag{2}$$

where G is the total number of similar patch groups; $\|C_i x\|_*$ is the nuclear norm of $C_i x$, taking a sum value of its singular values, namely, $\|C_i x\|_* = \sum_j |s_{i,j}|$. $s_i = [s_{i,1}, s_{i,2}, \ldots, s_{i,N}]$ denotes the singular value vector of $C_i x$, that is, $C_i x = U_i \Sigma_i V_i^T$, and s_i is a vector that contains the diagonal elements of Σ_i.

2.3. The Approximate Message Passing Algorithm. On the other hand, these minimizing problems (e.g., (1) and (2)) can be solved easily by iterative shrinkage/thresholding (IST) methods [26–28], alternating direction method of multipliers (ADMM) [29, 30], or Bregman iterative algorithms [31, 32]. The approximate message passing reconstruction algorithm defined by Donoho et al. [24] has recently become a popular algorithm for solving signal reconstruction problems in linear systems as defined in (1). It is based on the theory of belief propagation in graphical models. Unlike belief propagation that needs to calculate $2mn$ messages in each iteration, by employing quadratic approximation, the expression of each message can be simplified in the AMP algorithm, and the number of messages can be reduced to $m + n$ [33]. The final alternating expressions to solve the objective function $\min_x \|y - Ax\|_2^2/2 + \lambda\|\Psi(x)\|_1$ are

$$x^{(t+1)} = \eta\left(x^{(t)} + A^* z^{(t)}\right), \tag{3}$$

$$z^{(t)} = y - Ax^{(t)} + \frac{z^{(t-1)}\left\|\eta'\left(x^{(t-1)} + A^* z^{(t-1)}\right)\right\|_1}{m}, \tag{4}$$

where $\Psi(\cdot)$ denotes wavelet transforms; $x^{(t)}$ and $z^{(t)}$ are the estimates of x and the residual at iteration t. The iteration starts from $x^{(0)} = 0$ and $z^{(0)} = y$. A^* is the conjugate transpose of A. The functions $\eta(\cdot)$ and $\eta'(\cdot)$ are the wavelet threshold function and its first derivative, respectively. The last term in (4) is called the "Onsager reaction term" [24] in statistical physics. This Onsager reaction term helps improve the phase transition (trade-off between the measurement rate and signal sparsity) of the reconstruction process over existing IST algorithms [26–28]. We can summarize the AMP algorithm in three steps: a residue update step (i.e., (4)), a back-projection step to yield a noisy image $q^{(t)} = x^{(t)} + A^* z^{(t)}$,

Figure 1: The nonlocal similar patches in natural images.

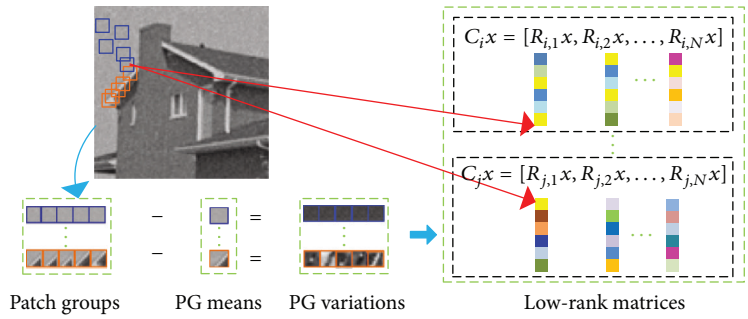


Figure 2: Illustration for the low-rank matrices construction.

and a proximal denoising correction (i.e., (3)). These steps are identical to the ones of IST methods. As a result, the AMP algorithm can be viewed as a special IST method. One benefit of this interpretation is that if $\mathfrak{R}(x)$ varies, only the proximal denoising operator needs to be altered, when the AMP algorithm is utilized to solve (1). Some AMP variants [13, 34, 35] have been proposed with various forms of $\mathfrak{R}(x)$, such as total variation [34], a Cauchy prior in the wavelet domain [35], and tree sparsity [13]. This motivates us to find a more suitable prior for natural images to improve the AMP algorithm.

3. Side Information-Aided LSM Prior Modeling

Image nonlocal self-similarity has been widely adopted in patch based CS image reconstruction methods. Despite the great success, most of the existing works exploit the nonlocal sparsity from the degraded image, which may cause the mismatching issue in the block matching stage. Moreover, they treat irregular and regular structures equally, resulting in an oversmoothed outcome. In fact, unlike large-scale edges, the fine-scale textures have much higher randomness in local structure and they are hard to characterize by using a local or nonlocal model. In this paper, we soften and complement the nonlocal sparsity for irregular structures by combining with local and global priors and propose a side information-aided LSM prior model for achieving detail-preserving and content-aware CS image reconstruction.

3.1. The Laplacian Scale Mixture Distribution. A random variable $\theta_i = \kappa_i^{-1}\varsigma_i$ is a Laplacian scale mixture if ς_i has a Laplacian distribution with scale 1; that is, $p(\varsigma_i) = \exp(-|\varsigma_i|)/2$, and the multiplier variable κ_i is a positive random variable with probability $p(\kappa_i)$ [21]. Supposing that ς_i and κ_i are independent, conditioned on the parameter κ_i, the coefficient θ_i has a Laplacian distribution with inverse scale $p(\theta_i \mid \kappa_i) = \kappa_i \exp(-\kappa_i|\theta_i|)/2$. The distribution over θ_i is therefore a continuous mixture of Laplacian distributions with different inverse scales:

$$p(\theta_i) = \int_0^\infty p(\theta_i \mid \kappa_i)\, p(\kappa_i)\, d\kappa_i = \int_0^\infty \frac{\kappa_i}{2} \exp(-\kappa_i |\theta_i|)\, p(\kappa_i)\, d\kappa_i. \tag{5}$$

The distribution in (5) is defined as a Laplacian scale mixture. Note that, for most choices of $p(\kappa_i)$, we do not have an analytical expression for $p(\theta_i)$.

3.2. The Proposed Model. In this section, we formulate the side information-aided LSM prior model and apply the MAP theory to estimate the original signal. As discussed earlier, the global information of image refers to a prediction that is actually a reference image. We use this reference image denoted by u as side information which is supposed to be known in each iteration to assist the reconstruction.

Using Bayesian formula, we might derive the following MAP estimation problem:

$$\begin{aligned}\widehat{x} &= \arg\min_x p(x \mid y,u)\\ &= \arg\min_x \frac{p(y,u \mid x)\,p(x)}{p(y,u)}\\ &= \arg\min_x \frac{p(y \mid x)\,p(u \mid x)\,p(x)}{p(y,u)}\\ &= \arg\min_x p(y \mid x)\,p(u \mid x)\,p(x).\end{aligned} \tag{6}$$

The side information u and the measurements y are supposed to be independent. According to (6), we need to define three terms $p(y \mid x)$, $p(u \mid x)$, and $p(x)$ for the MAP estimation.

First, the additive noise w is assumed to be white Gaussian with variance σ^2, that is, $w \sim N(0,\sigma^2)$. Thus, we have the following likelihood of CS:

$$p(y \mid x) = \frac{1}{(2\pi\sigma^2)^{m/2}} \exp\left(-\frac{1}{2\sigma^2}\left\|y - Ax\right\|_2^2\right). \tag{7}$$

Second, we characterize the nonlocal sparsity of image with the Laplacian scale mixture model. Recall that, in Section 2, we define $s_i = [s_{i,1}, s_{i,2}, \ldots, s_{i,N}]$ as the singular value vector of the low-rank matrix $C_i x$. For each coefficient $s_{i,j}$, we assign it a Laplacian distribution with a variable scale parameter,

$$\begin{aligned}p(s_{i,j}) &= \int_0^\infty p(s_{i,j} \mid \gamma_{i,j})\,p(\gamma_{i,j})\,d\gamma_{i,j}\\ &= \int_0^\infty \frac{\gamma_{i,j}}{2} \exp\left(-\gamma_{i,j}\left|s_{i,j}\right|\right) p(\gamma_{i,j})\,d\gamma_{i,j},\end{aligned} \tag{8}$$

with a Gamma distribution prior over the scale parameter, that is, $\gamma_{i,j} \sim \text{Gamma}(a,b)$. Note that the mean of the Laplacian distribution is 0. Hence, the mean subtraction should be carried out as shown in Figure 2. The observation that the singular values can be modeled by a Laplacian distribution has been proposed and validated in [36]. Here, we extend this idea by viewing scale parameters as random variables for achieving a better spatial adaptation. Assume that the sparse coefficients are i.i.d., and then the LSM prior of the coefficients can be expressed as $p(s) = \prod_{i=1}^{G}\prod_{j=1}^{N} p(s_{i,j})$.

Third, for irregular structures, the packed patches are less similar. We wish to preserve the relative difference among these patches. Supposing that the reference image u is available, we wish to get a solution that has a small Euclidean distance from u while being constrained by the self-similarity prior:

$$\begin{aligned}p(u \mid x) &= \prod_{i=1}^{H} p(u_i \mid x_i)\\ &= \prod_{i=1}^{H} \exp\left(-\frac{1}{2\sigma_u^2}\left\|x_i - u_i\right\|_2^2\right),\end{aligned} \tag{9}$$

where u_i and x_i denote the patches located at the ith position and H is the total number of patches. In practice, the reference image u is produced by minimizing the data fidelity term with the gradient descent method [26]. Note that the resulting image u has fine details as well as a small amount of noise; thus the noise level might be a little higher than the estimated one. The image u which integrates the outcomes at the independent processing stage can represent the global information of image. This approach is similar to the one reported in [37], which is a denoising method. It reduces the local-global gap by encouraging the overlapping patches to reach an agreement before they merge their forces by the averaging.

At last, when the packed patches have an obvious dissimilarity, a local constraint imposed on each local patch is added by enforcing a small (i.e., sparse) prediction error $e_i = x_i - v_i$, where v is the noiseless version of the image u. To get v, the mean filtering is applied. The singular value vector of the prediction error of the ith patch is denoted by $r_i = [r_{i,1}, r_{i,2}, \ldots, r_{i,K}]$, which is imposed on a LSM distribution $p(r_i) = \prod_{j=1}^{K} p(r_{i,j})$, where

$$\begin{aligned}p(r_{i,j}) &= \int_0^\infty p(r_{i,j} \mid \pi_{i,j})\,p(\pi_{i,j})\,d\pi_{i,j}\\ &= \int_0^\infty \frac{\pi_{i,j}}{2} \exp\left(-\pi_{i,j}\left|r_{i,j}\right|\right) p(\pi_{i,j})\,d\pi_{i,j},\end{aligned} \tag{10}$$

with a Gamma distribution prior over the scale parameter, that is, $\pi_{i,j} \sim \text{Gamma}(c,d)$. K is the total number of singular values of the ith patch.

The proposed hierarchical model can be summarized in Figure 3. We will have an objective function that can be maximized with respect to x, if we observe the latent variable γ and π. The standard approach in machine learning when confronted with such a problem is the EM algorithm. Note that once the spare coefficients (i.e., s and r) are determined, the image x can be obtained by averaging all reconstructed patches.

4. EM-AMP Algorithm for CS Recovery

In this section, we simultaneously learn the hidden parameters and do inference. To accomplish this task, we embed the AMP algorithm within an EM framework.

4.1. EM Learning of the Prior Parameters. We use Jensen's inequality and obtain the following upper bound on the posterior likelihood, denoted by $J(\Theta, x)$, where $\Theta = \{Q, D\}$, according to (6):

$$\begin{aligned}-\log p(x \mid y,u) \le &-\log p(y \mid x) - \log p(u \mid x)\\ &- \int_\gamma Q(\gamma) \log \frac{p(s,\gamma)}{Q(\gamma)}\,dr\\ &- \int_\pi D(\pi) \log \frac{p(r,\pi)}{D(\pi)}\,d\pi,\end{aligned} \tag{11}$$

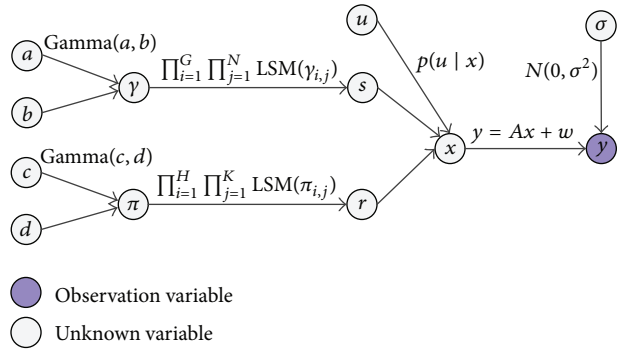


Figure 3: The proposed hierarchical model.

where

$$\begin{aligned}\log p(s,\gamma) &= \log p(s \mid \gamma) + \log p(\gamma)\\ &= \sum_{i=1}^{G}\sum_{j=1}^{N}\left(-\gamma_{i,j}\left|s_{i,j}\right| + \log\frac{\gamma_{i,j}}{2}\right)\\ &\quad + \log p(\gamma),\\ \log p(r,\pi) &= \log p(r \mid \pi) + \log p(\pi)\\ &= \sum_{i=1}^{H}\sum_{j=1}^{K}\left(-\pi_{i,j}\left|r_{i,j}\right| + \log\frac{\pi_{i,j}}{2}\right)\\ &\quad + \log p(\pi).\end{aligned} \tag{12}$$

Performing coordinate descent in the auxiliary function $J(\Theta, x)$ leads to the following updates that are usually called the E step and the M step:

$$E\ step\text{: }\Theta^{(t+1)} = \arg\min_{\Theta} J\left(\Theta, x^{(t)}\right), \tag{13}$$

$$M\ step\text{: }x^{(t+1)} = \arg\min_{x} J\left(\Theta^{(t+1)}, x\right). \tag{14}$$

The E step and the M step represent the prior learning and the signal reconstruction, respectively. First, we have equality in the Jensen inequality if $Q(\gamma) = p(\gamma \mid s)$ and $D(\pi) = p(\pi \mid r)$, which implies that the E step can be reduced to $Q^{(t+1)} = p(\gamma \mid s^{(t)})$ and $D^{(t+1)} = p(\pi \mid r^{(t)})$. Second, let τ and ε denote the expectation with respect to $Q(\gamma)$ and $D(\pi)$, respectively; that is, $\tau_{i,j} = \int_{\gamma_{i,j}} \gamma_{i,j} Q(\gamma_{i,j}) d\gamma_{i,j}$ and $\varepsilon_{i,j} = \int_{\pi_{i,j}} \pi_{i,j} D(\pi_{i,j}) d\pi_{i,j}$. The M Step (14) simplifies to

$$\begin{aligned}x^{(t+1)} = \arg\min_{x,s,r} &\frac{\left\|y - Ax\right\|_2^2}{2\sigma^2} + \sum_{i=1}^{H}\frac{\left\|x_i - u_i\right\|_2^2}{2\sigma_u^2}\\ &+ \sum_{i=1}^{G}\sum_{j=1}^{N}\tau_{i,j}^{(t+1)}\left|s_{i,j}\right| + \sum_{i=1}^{H}\sum_{j=1}^{K}\varepsilon_{i,j}^{(t+1)}\left|r_{i,j}\right|.\end{aligned} \tag{15}$$

Finally, we have the proposed objective function for CS recovery. These four terms are the data fidelity term, side information-aided term, the nonlocal regularization, and the local regularization, respectively.

The Gamma distribution and Laplacian distribution are conjugate; that is, the posterior probability of $\gamma_{i,j}$ (or $\pi_{i,j}$) given $s_{i,j}$ (or $r_{i,j}$) is also a Gamma distribution with parameters $a + 1$ (or $c + 1$) and $b + |s_{i,j}|$ (or $d + |r_{i,j}|$). Hence, the expectations are given by

$$\begin{aligned}\tau_{i,j}^{(t+1)} &= \int_{\gamma_{i,j}} \gamma_{i,j} p\left(\gamma_{i,j} \mid s_{i,j}^{(t)}\right) d\gamma_{i,j} = \frac{(a+1)}{\left(b + \left|s_{i,j}^{(t)}\right|\right)},\\ \varepsilon_{i,j}^{(t+1)} &= \int_{\pi_{i,j}} \pi_{i,j} p\left(\pi_{i,j} \mid r_{i,j}^{(t)}\right) d\pi_{i,j} = \frac{(c+1)}{\left(d + \left|r_{i,j}^{(t)}\right|\right)}.\end{aligned} \tag{16}$$

4.2. Image Recovery via the AMP Algorithm. Once the prior parameters are estimated, the composite regularization problem, that is, (15), can be solved with various CS reconstruction algorithms [29–32]. Owing to its promising performance, the AMP algorithm [24, 25] is employed. As discussed earlier, the proximal operator varies according to the regularization term $\mathfrak{R}(x)$. Hence we have the following proximal operator $\ell(q^{(t)})$ on the basis of (15):

$$\begin{aligned}x^{(t+1)} = \arg\min_{x,s,r} &\frac{\left\|x - q^{(t)}\right\|_2^2}{2\sigma^2} + \sum_{i=1}^{H}\frac{\left\|x_i - u_i\right\|_2^2}{2\sigma_u^2}\\ &+ \sum_{i=1}^{G}\sum_{j=1}^{N}\tau_{i,j}^{(t+1)}\left|s_{i,j}\right| + \sum_{i=1}^{H}\sum_{j=1}^{K}\varepsilon_{i,j}^{(t+1)}\left|r_{i,j}\right|,\end{aligned} \tag{17}$$

where $q^{(t)} = x^{(t)} + A^* z^{(t)}$ denotes a noisy image yielded by back-projection and $\ell(q^{(t)})$ is a denoising operation. To solve this composite regularization problem, that is, (17), we decompose it into two simpler regularization subproblems by using the composite splitting algorithm (CSA) [7], which includes three steps: (1) splitting variable x into two variables x_1 and x_2, (2) performing operator splitting to minimize the nonlocal regularization and the local regularization subproblems over x_1 and x_2, respectively, and (3) obtaining the solution x by linear combination of x_1 and x_2, that is, $x = \omega_1 x_1 + \omega_2 x_2$. Subproblems can be written as

$$x_1^{(t+1)} = \arg\min_{x,s} \frac{\left\|x - q^{(t)}\right\|_2^2}{2\sigma^2} + \sum_{i=1}^{H}\frac{\left\|x_i - u_i\right\|_2^2}{2\sigma_u^2}$$

$$+\sum_{i=1}^{G}\sum_{j=1}^{N}\tau_{i,j}^{(t+1)}\left|s_{i,j}\right| = \arg\min_{x,s}\frac{\left\|x-\left(q^{(t)}+\sigma^2u^{(t)}/\sigma_u^2\right)\right\|_2^2}{2\sigma^2} + \sum_{i=1}^{G}\sum_{j=1}^{N}\tau_{i,j}^{(t+1)}\left|s_{i,j}\right|, \tag{18}$$

$$x_2^{(t+1)} = \arg\min_{x,r}\frac{\left\|x-\left(q^{(t)}+\sigma^2u^{(t)}/\sigma_u^2\right)\right\|_2^2}{2\sigma^2} + \sum_{i=1}^{H}\sum_{j=1}^{K}\varepsilon_{i,j}^{(t+1)}\left|r_{i,j}\right|. \tag{19}$$

In order to solve (18), we transform the data from spatial domain to SVD domain for the unity of variable representation by (1) dividing the images into overlapped patches; (2) building the low-rank matrices $X_i = C_i x$ and $Q_i^{(t)} = C_i(q^{(t)} + \sigma^2 u^{(t)}/\sigma_u^2)$, and then we have $\|x - (q^{(t)} + \sigma^2u^{(t)}/\sigma_u^2)\|_2^2 = \sum_{i=1}^{G}\|X_i - Q_i^{(t)}\|_F^2/\alpha$, where α is a scale factor for balancing the increasing energy caused by the repetitive computation because of the overlap division and $\|\cdot\|_F$ is the Frobenius norm [17]; (3) performing the singular value decomposition on X_i and $Q_i^{(t)}$ to get the singular value vectors $s_i = [s_{i,1}, s_{i,2}, \ldots, s_{i,N}]$ and $\rho_i = [\rho_{i,1}, \rho_{i,2}, \ldots, \rho_{i,N}]$. Based on the von Neumann trace inequality [38], we know that $\mathrm{tr}((Q_i^{(t)})^T X_i)$ achieves its upper bound $\sum_{j=1}^{N}\rho_j s_j$. Then we have

$$\begin{aligned}\left\|X_i - Q_i^{(t)}\right\|_F^2 &= \mathrm{tr}\left(\left(Q_i^{(t)}\right)^T Q_i^{(t)}\right) - 2\mathrm{tr}\left(\left(Q_i^{(t)}\right)^T X_i\right) + \mathrm{tr}\left((X_i)^T X_i\right)\\ &\geq \sum_{j=1}^{N}\rho_j^2 - 2\sum_{j=1}^{N}\rho_j s_j + \sum_{j=1}^{N}s_j^2.\end{aligned} \tag{20}$$

By incorporating (20) into (18), it yields

$$s^{(t+1)} = \arg\min_{s}\sum_{i=1}^{G}\left\{\frac{\sum_{j=1}^{N}\rho_{i,j}^2 - 2\sum_{j=1}^{N}\rho_{i,j}s_{i,j} + \sum_{j=1}^{N}s_{i,j}^2}{2\alpha\sigma^2} + \sum_{j=1}^{N}\tau_{i,j}^{(t+1)}\left|s_{i,j}\right|\right\} = \arg\min_{s}\sum_{i=1}^{G}\sum_{j=1}^{N}\left(s_{i,j} - \rho_{i,j} + \alpha\sigma^2\tau_{i,j}^{(t+1)}\right)^2. \tag{21}$$

Taking a derivative with respect to $s_{i,j}$, we have

$$s_{i,j}^{(t+1)} = \max\left(\rho_{i,j} - \alpha\sigma^2\tau_{i,j}^{(t+1)}, 0\right), \tag{22}$$

which is the global optimum of $s_{i,j}$. The noise variance σ^2 is obtained by maximum likelihood estimation [39]; that is, $\sigma^2 = \|z^{(t)}\|_2^2/m$. Afterward, we obtain the matrix constructed by similar patches, that is, $X_i^{(t+1)} = U_i\,\mathrm{diag}(s_i^{(t+1)})V_i^T$, and then recover $x_1^{(t+1)}$ by averaging all reconstructed patches.

Similarly, we can derive the global optimum of another subproblem by incorporating (20) into (19):

$$r_{i,j}^{(t+1)} = \max\left(\phi_{i,j} - \sigma^2\pi_{i,j}^{(t+1)}, 0\right), \tag{23}$$

where ϕ_i is the singular value vector of the ith patch. Then $x_2^{(t+1)}$ can be recovered by aggregating all reconstructed patches. Because r is the singular value of the prediction error, the reconstructed image solved by this subproblem is the prediction error in fact. Thus, we add v to obtain the final image $x_2^{(t+1)}$.

To apply the AMP algorithm, computing the Onsager term $z^{(t)}\|\ell'(q^{(t)})\|_1/m$, which involves computing the derivative $\ell'(q^{(t)})$, is required. It is not easy to get $\ell'(q^{(t)})$, since $\ell(q^{(t)})$ do not have an explicit input-output relation. Thanks to the Monte Carlo (MC) method [40], we can simulate $\ell'(q^{(t)})$ with random numbers. This method has been used to estimate the derivative of BM3D denoiser in [18]. More details can be found in [18].

4.3. Content-Aware Strategy. In order to preserve fine details, we distinguish irregular structures from regular structures through the similarity of packed patches. The local regularization and side information-aided term are only utilized for irregular structures. They are adopted to maintain differences among packed patches. We use the normalized mean squared error (NMSE) as the similarity measure:

$$E = \sum_{j}\frac{\left\|\widetilde{x}_i - x_j\right\|_2^2}{\left\|\widetilde{x}_i\right\|_2^2}, \tag{24}$$

where $\widetilde{x}_i$ is an exemplar patch located at the ith position and x_j denotes the jth patch in the similar patch group. The objective function (17) is then modified to

$$x^{(t+1)} = \arg\min_{x,s,r}\frac{\left\|x-q^{(t)}\right\|_2^2}{2\sigma^2} + \sum_{i=1}^{H}W_i\odot\frac{\left\|x_i-u_i\right\|_2^2}{2\sigma_u^2} + \sum_{i=1}^{G}\sum_{j=1}^{N}\tau_{i,j}^{(t+1)}\left|s_{i,j}\right| + \sum_{i=1}^{H}\sum_{j=1}^{K}\overline{W}_i\odot\varepsilon_{i,j}^{(t+1)}\left|r_{i,j}\right|, \tag{25}$$

where W_i and $\overline{W}_i$ denote the weights that correspond to the ith patch and $\odot$ is the Hadamard product. W_i and $\overline{W}_i$ are determined by

$$W_i = \begin{cases}0, & \text{if } E < \mathrm{TH1}\\ 1, & \text{otherwise,}\end{cases}$$

Input: $y, A, T, a, b, c, d, \omega_1, \omega_2$, TH1, TH2, $x^0 = 0$, $z^0 = y$.
for $t = 0$ **to** $T - 1$ **do**
(a) Approximate the Onsager correction term via MC.
(b) Update the residual $z^{(t)} = y - Ax^{(t)} + z^{(t-1)}\|\ell'(q^{(t-1)})\|_1/m$.
(c) Obtain the noisy image $q^{(t)} = x^{(t)} + A^* z^{(t)}$.
(d) Calculate the proximal operator $\ell(q^{(t)})$ (i.e., solve (25))
for $i = 1$ **to** G **do**
(I) Construct the low-rank matrix $Q_i^{(t)}$.
(II) Distinguish irregular structures from regular structures with the similarity measure E, and set W_i and $\overline{W}_i$ via (26).
(III) Perform the SVD on $Q_i^{(t)}$ to get the singular value vector $[\rho_{i,1}, \rho_{i,2}, \ldots, \rho_{i,N}]$.
(IV) Estimate the expectations of scale parameters $\tau_{i,j}^{(t+1)}$ and $\varepsilon_{i,j}^{(t+1)}$ via (16), and the noise variance $\sigma^2 = \|z^{(t)}\|_2^2/m$.
(V) Compute the global optimums of coefficients $s_{i,j}^{(t+1)}$ and $r_{i,j}^{(t+1)}$ via (22) and (23).
(VI) If $i = G$, recover the whole image $x^{(t+1)} = \omega_1 x_1^{(t+1)} + \omega_2 x_2^{(t+1)}$ by aggregating all recovered pixels.
end for
end for

Algorithm 1: CS image recovery via SI-LSM-AMP.

$$\overline{W}_i = \begin{cases} 1, & \text{if } E > \text{TH2} \\ 0, & \text{otherwise,} \end{cases} \tag{26}$$

where TH1 and TH2 (TH2 > TH1) are two threshold parameters. First, if the patches are similar, that is, $E < T$H1, only the nonlocal regularization is utilized. Second, if the packed patches are not so similar, that is, TH1 $< E <$ TH2, the side information-aided term is added. At last, when the patches have an obvious dissimilarity, that is, $E >$ TH2, the local regularization is superadded.

In general, the proposed algorithm is summarized in Algorithm 1, named as the AMP algorithm with side information-aided LSM priors (SI-LSM-AMP). We also propose an algorithm as a by-product, named as the AMP algorithm with nonlocal LSM prior (NL-LSM-AMP), which is a special form of SI-LSM-AMP by setting $E <$ TH1 for all image patches. By comparing SI-LSM-AMP with NL-LSM-AMP in the experiments, one can validate the effectiveness of our content-aware strategy.

5. Experiments

In this section, we report the experimental results of the proposed CS recovery method SI-LSM-AMP.

5.1. Experiment Setup. We generate the CS measurements by randomly sampling the Fourier transform coefficients of test images; that is, A is partial Fourier transform with m rows and n columns. Thus, the sampling ratio is m/n. We follow the sampling strategy of previous works ([7, 12]), which randomly choose more Fourier coefficients from low frequency and less on high frequency and set the sampling ratio near to 0.2, as CS imaging is always interested in low sampling ratio cases. All measurements are mixed with Gaussian white noise with standard deviations 5 and 15, representing the environments with low noise and high noise, respectively. Peak Signal-to-Noise Ratio (PSNR) is used for quantitative evaluation.

We conduct experiments on 22 natural images, as shown in Figure 4. Besides the NL-LSM-AMP algorithm, we compare the SI-LSM-AMP algorithm with image reconstruction algorithms, including the original AMP method [24], two tree-based algorithms: WaTMRI [12] and Turbo-AMP [13], and two nonlocal sparsity algorithms: NLR-CS [17] and BM3D-AMP [19]. For fair comparisons, all codes are downloaded from the authors' websites. All algorithms terminate after 50 iterations, except 20 iterations for Turbo-AMP and 260 iterations for NLR-CS. For the WaTMRI algorithm, in order to achieve the best result, the regularization parameters are tuned to 0.8 and 0.35. The original AMP uses Daubechies wavelet to decompose the images. Then the wavelet denoising is applied with a threshold 0.8σ. For the rest of algorithms, default settings in their codes are adopted. All experiments are on a desktop with 3.80 GHz AMD A10-5800K CPU. Matlab version is R2014a.

The main parameters of the proposed algorithms are set as follows: patch size is 6×6; a total of 36 similar patches are selected for each exemplar patch. To reduce the computational complexity, we extract exemplar image patch in every 5 pixels along both horizontal and vertical directions. For better CS recovery performance, some parameters are tuned empirically, including (1) $\sigma_u^2 = 3.3\sigma^2$, which implies that the proportion between the data fidelity term and the side information-aided term is 1 : 0.3; (2) $\alpha = 5.2$ in (22); (3) the parameters of Gamma distribution $a = c = 0$ and $b = d = 0.01$ as suggested in [21]; (4) the combination parameters $\omega_1 = 2/3$ and $\omega_2 = 1/3$. Recall that the noise level should be a little higher (1.08 times in our experiments) than the estimated one because the reference image introduces a small amount of noise. The experimental results including average PSNR, visual quality, and runtime are present.

Figure 4: The 22 test images.

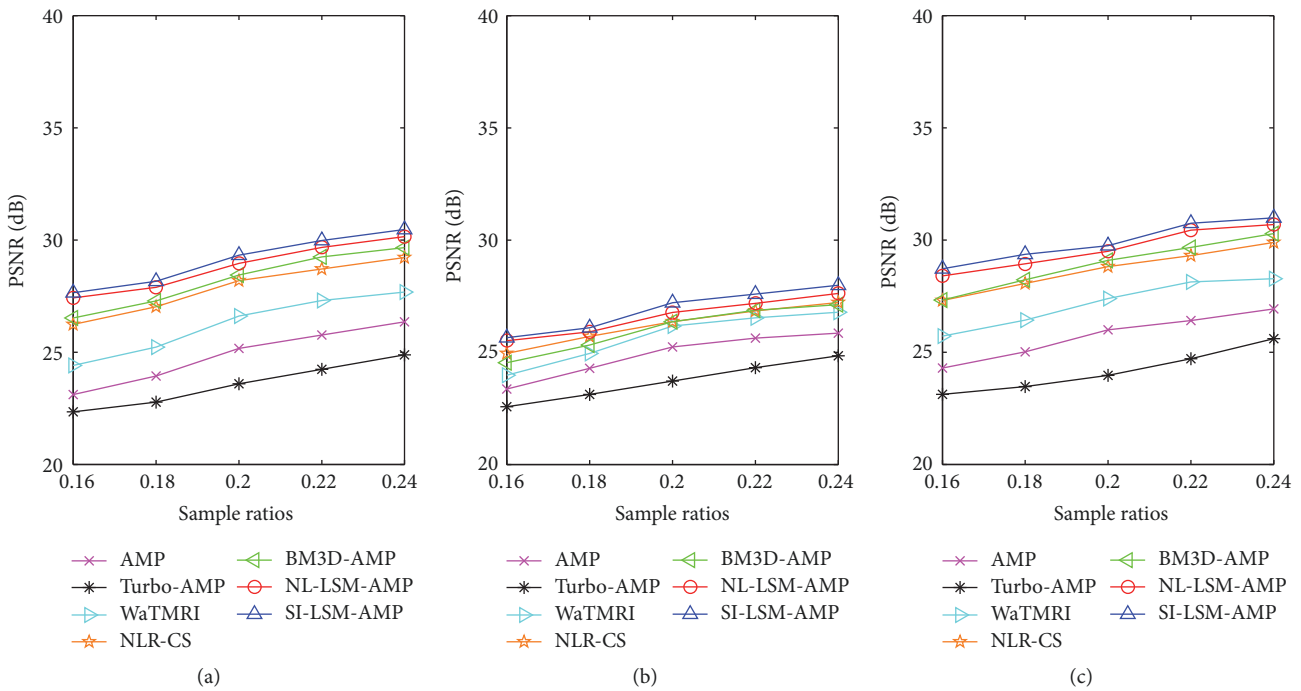


Figure 5: Average PSNR at different sampling ratios with measurement noise with standard deviation 5. (a) Comparisons on 22 natural images; (b) comparisons on Baboon image; (c) comparisons on Boat image.

5.2. Average PSNR Evaluation. To reduce the randomness, we run each experiment 5 times for each image. The average PSNR results are shown in Figures 5 and 6. Figure 5 shows the PSNR results in the environment with low noise, while Figure 6 shows the ones with high noise. From these figures, one can draw the following four conclusions. First, four nonlocal sparsity-based methods SI-LSM-AMP, NL-LSM-AMP, BM3D-AMP, and NLR-CS perform better than

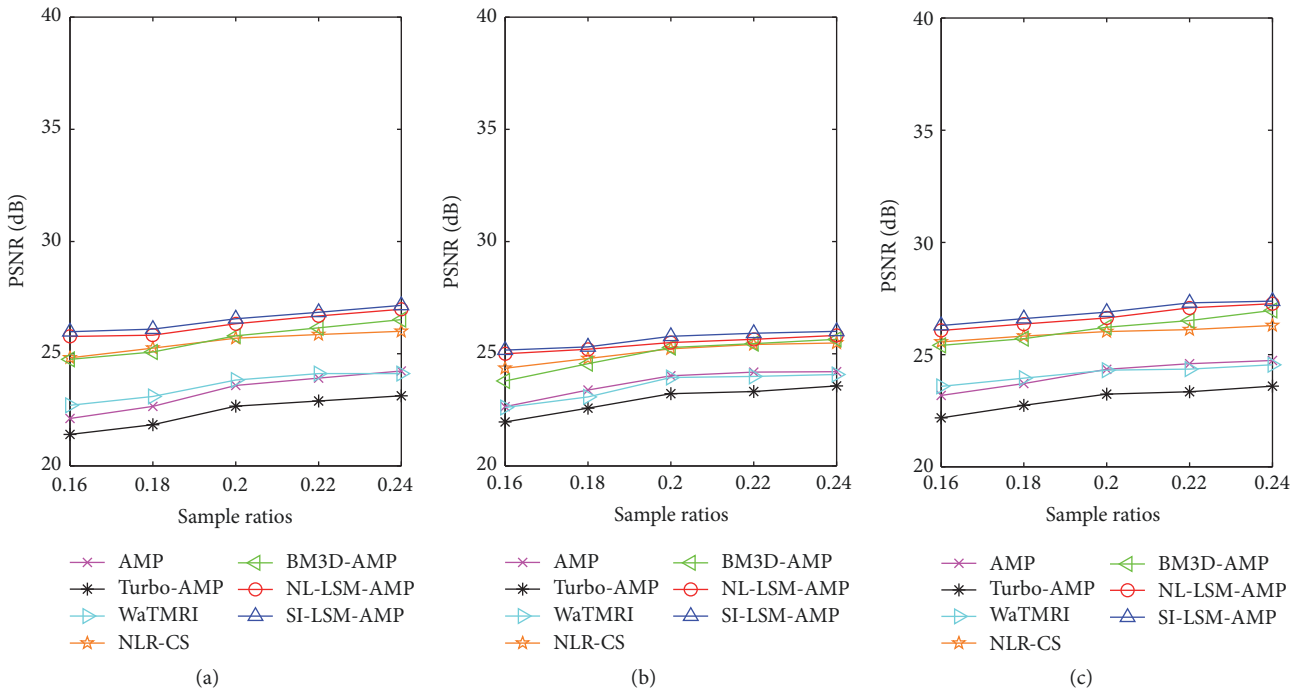
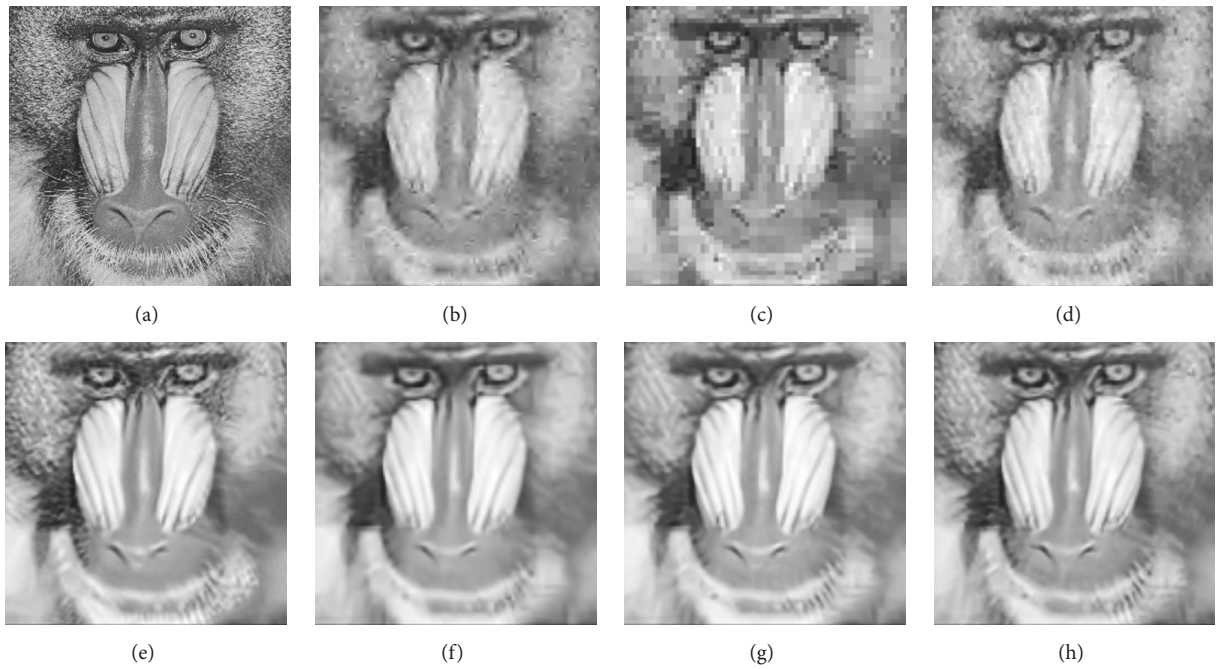


Figure 6: Average PSNR at different sampling ratios with measurement noise with standard deviation 15. (a) Comparisons on 22 natural images; (b) comparisons on Baboon image; (c) comparisons on Boat image.

others, which implies that the assumption of the nonlocal sparsity structure is more appropriate for natural images when compared to the tree structure and the standard sparsity. Second, the highest PSNR results are achieved by the proposed algorithm SI-LSM-AMP. In fact, the PSNR gain of the SI-LSM-AMP algorithm over the next-best algorithm NL-LSM-AMP can be as much as 0.25 dB on average. This result validates the effectiveness of our content-aware strategy that treats irregular and regular structures differently. Third, through the use of the LSM prior model, the SI-LSM-AMP algorithm and the NL-LSM-AMP algorithm outperform the other nonlocal sparsity-based AMP method BM3D-AMP (in fact, by 0.88 dB and 0.63 dB on average). Therefore, we can conclude that the LSM prior model is more appropriate for representing the nonlocal sparsity of natural images. At last, the PSNR curves decline with the decrease in sampling ratio or with the increase in measurement noise. The SI-LSM-AMP algorithm always performs better with low sampling ratio or strong measurement noise. For better quality perception, experiments performed on Baboon and Boat images present the similar results in Figures 5(b) and 5(c) and Figures 6(b) and 6(c), respectively. In a word, these results validate the superiority of the proposed SI-LSM-AMP algorithm in objective quality and the effectiveness of the LSM prior model.

5.3. Visual Quality and Runtime Evaluation in Lower-Power Noise Environment. Figures 7 and 8 show the visual comparisons of the reconstructed results on Baboon and Boat images with 20% sampling by different methods with measurement noise with standard deviation 5, while the corresponding iterative curves are given in Figures 9 and 10, respectively. From Figures 7 and 8, we can clearly see that four nonlocal sparsity-based methods are still better than others. Among them, the SI-LSM-AMP algorithm enjoys great advantages over the NL-LSM-AMP algorithm in producing clearer image, for example, on the area of hair (Figure 7) and beach (Figure 8). It can not only perfectly reconstruct large-scale sharp edges but also well recover small-scale fine structures. The images reconstructed by the NL-LSM-AMP algorithm and the BM3D-AMP algorithm are too smooth. These two methods all have a strong assumption of the nonlocal self-similarity structure. However, many images with irregular structures do not strictly follow this assumption. We could find that the SI-LSM-AMP algorithm has great superiority on the images with irregular structures, because we combine local and global sparsity, which soften and complement the nonlocal self-similarity structure assumption for irregular structures. We also compute the PSNR as well as the structural similarity index (SSIM), which better reflects the visual quality of the images. The SI-LSM-AMP algorithm achieves the highest quantitative results. The superiority of the proposed SI-LSM-AMP in visual quality could be demonstrated by these results.

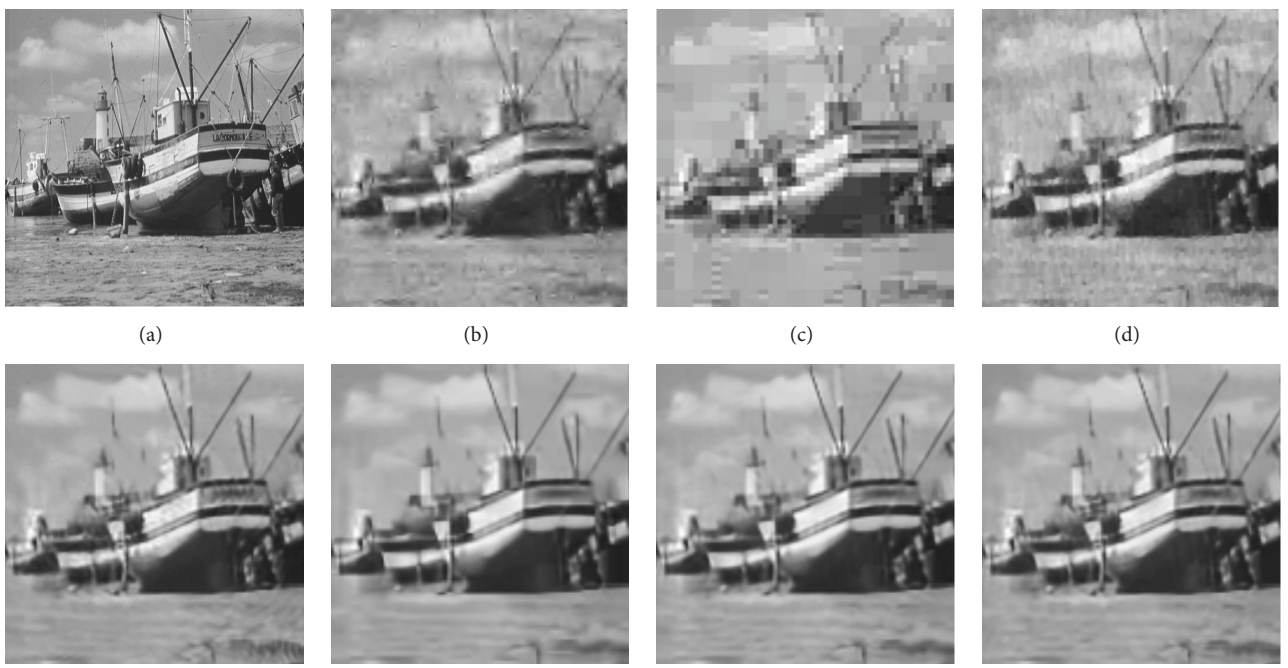
The CPU time and PSNR are traced in each iteration for each of the methods. Figures 9 and 10 present the CPU time versus PSNR curves and iteration number versus PSNR curves, respectively. We can see that the SI-LSM-AMP algorithm achieves higher PSNR results after about 30 iterations in Figures 9(a) and 10(a) and after about 100 seconds in Figures 9(b) and 10(b). Note that, considering the fact that it is hard to distinguish irregular structures from



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

(e) (f) (g) (h)

Figure 7: Visual comparisons on Baboon image at 20% sampling ratio with measurement noise with standard deviation 5. (a) The original image; (b) AMP (PSNR = 25.44 dB, SSIM = 0.6968); (c) Turbo-AMP (PSNR = 23.82 dB, SSIM = 0.6103); (d) WaTMRI (PSNR = 26.37 dB, SSIM = 0.7590); (e) NLR-CS (PSNR = 26.56 dB, SSIM = 0.7877); (f) BM3D-AMP (PSNR = 26.56 dB, SSIM = 0.7478); (g) NL-LSM-AMP (PSNR = 26.97 dB, SSIM = 0.7792); (h) SI-LSM-AMP (PSNR = 27.41 dB, SSIM = 0.8051).



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

(e) (f) (g) (h)

Figure 8: Visual comparisons on Boat image at 20% sampling ratio with measurement noise with standard deviation 5. (a) The original image; (b) AMP (PSNR = 26.00 dB, SSIM = 0.7702); (c) Turbo-AMP (PSNR = 23.89 dB, SSIM = 0.6748); (d) WaTMRI (PSNR = 27.39 dB, SSIM = 0.8192); (e) NLR-CS (PSNR = 28.81 dB, SSIM = 0.8687); (f) BM3D-AMP (PSNR = 29.09 dB, SSIM = 0.8755); (g) NL-LSM-AMP (PSNR = 29.42 dB, SSIM = 0.8861); (h) SI-LSM-AMP (PSNR = 29.67 dB, SSIM = 0.8927).

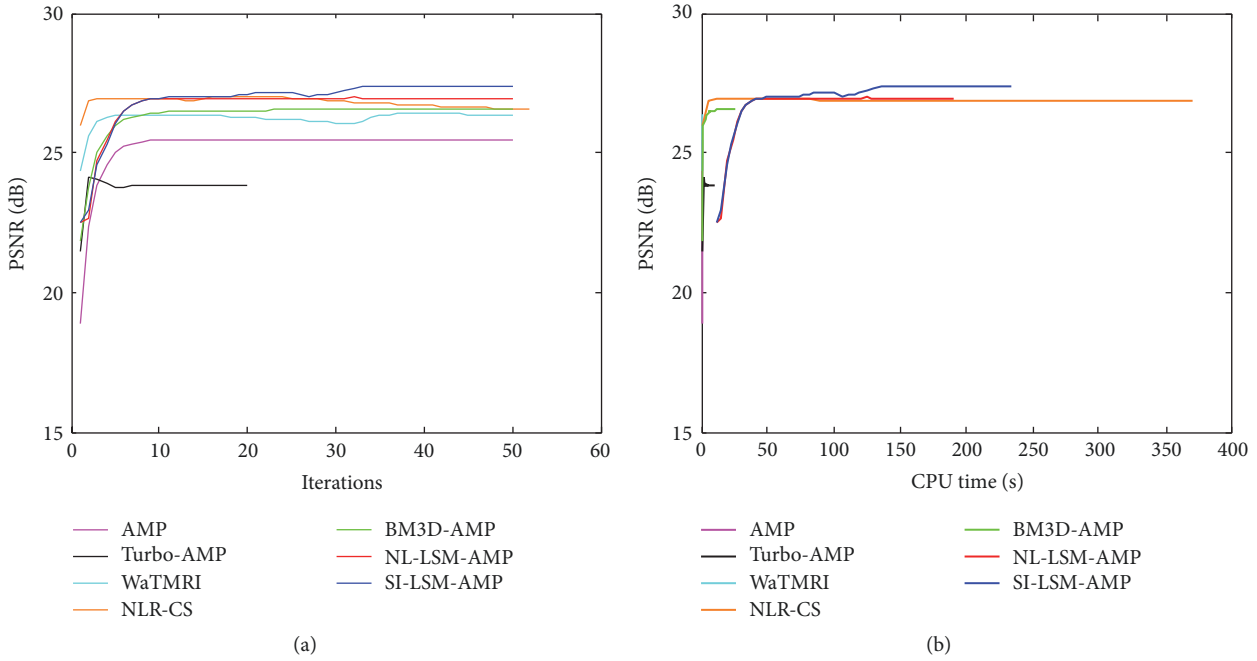


Figure 9: Iterative curves on Baboon image at 20% sampling ratio with measurement noise with standard deviation 5. (a) Average PSNR to iterations; (b) average PSNR to CPU time.

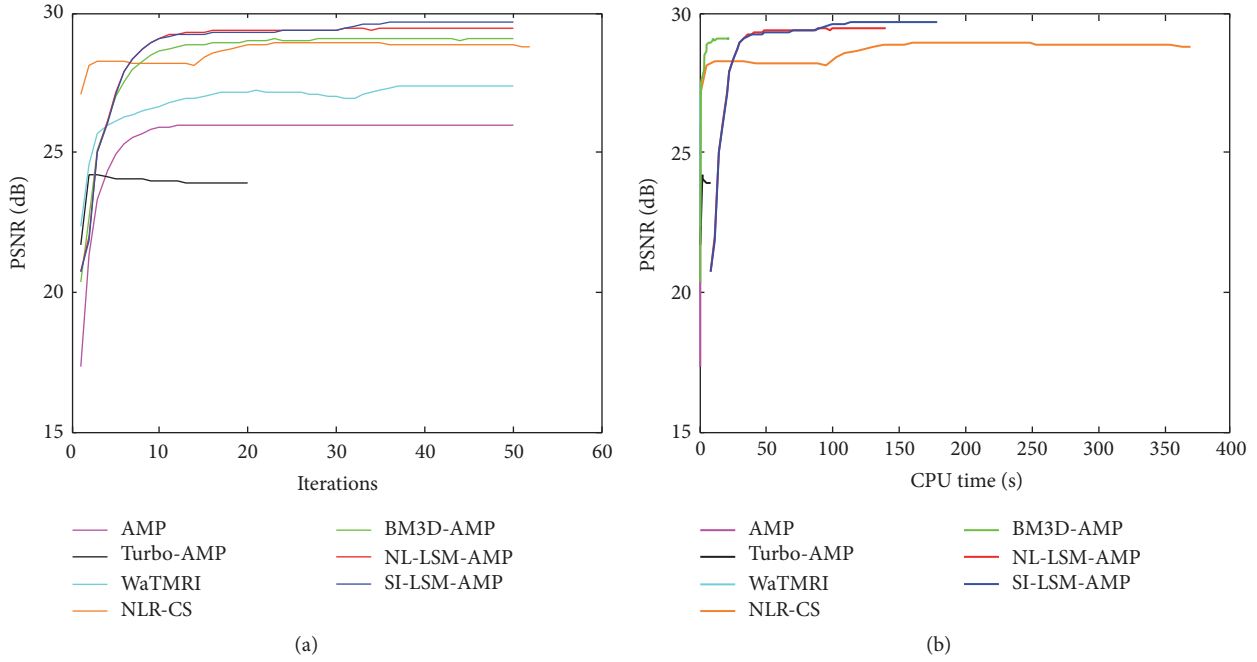


Figure 10: Iterative curves on Boat image at 20% sampling ratio with measurement noise with standard deviation 5. (a) Average PSNR to iterations; (b) average PSNR to CPU time.

regular structures at the beginning stage, we only use the nonlocal sparsity constraint to obtain an initial reconstructed image in the first 30 iterations of the SI-LSM-AMP algorithm. Therefore, the PSNR results of the SI-LSM-AMP algorithm in the first 30 iterations are identical to the ones of the NL-LSM-AMP algorithm. The SI-LSM-AMP algorithm is relatively slow. The main computational burdens are introduced by iteratively applying SVD on each of the patch groups. The computational time of our algorithm can be further reduced through the use of the parallel computing method or C language development. Besides, in the case of NLR-CS with more iterations, its results are every five iterations presented

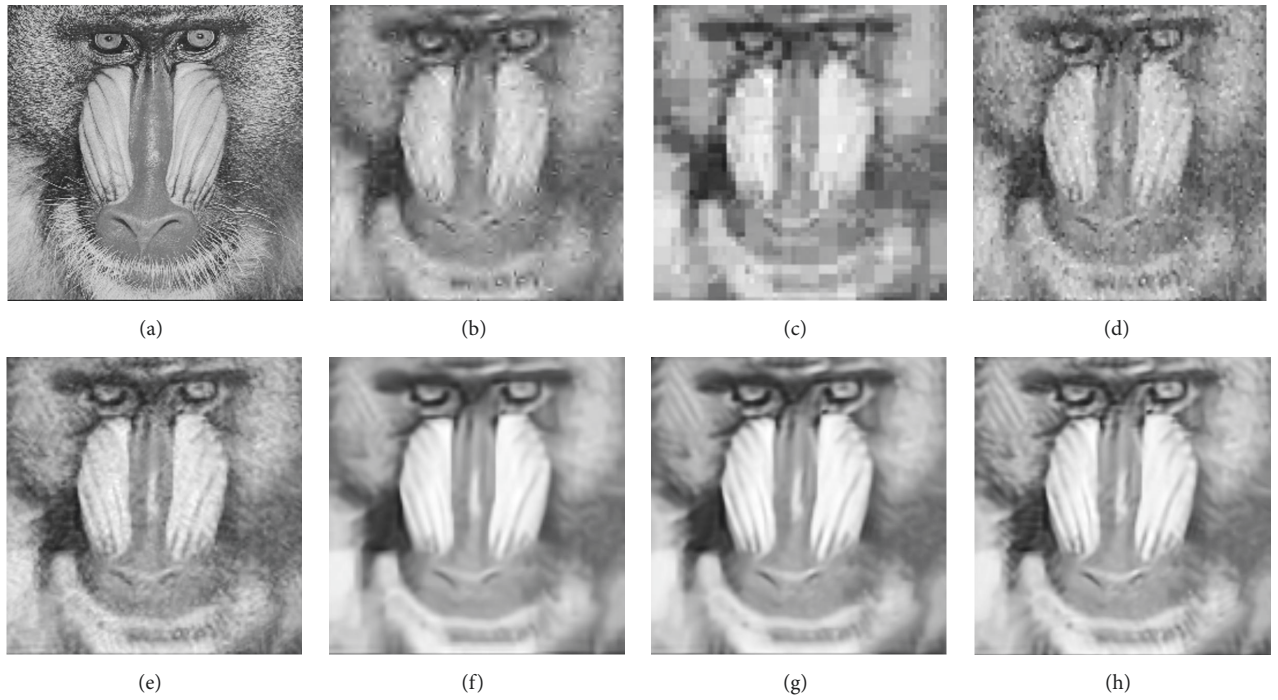


Figure 11: Visual comparisons on Baboon image at 20% sampling ratio with measurement noise with standard deviation 15. (a) The original image; (b) AMP (PSNR = 24.16 dB, SSIM = 0.6047); (c) Turbo-AMP (PSNR = 23.27 dB, SSIM = 0.5405); (d) WaTMRI (PSNR = 23.99 dB, SSIM = 0.6342); (e) NLR-CS (PSNR = 25.33 dB, SSIM = 0.6972); (f) BM3D-AMP (PSNR = 25.32 dB, SSIM = 0.6755); (g) NL-LSM-AMP (PSNR = 25.55 dB, SSIM = 0.6962); (h) SI-LSM-AMP (PSNR = 25.82 dB, SSIM = 0.7159).

in Figures 9 and 10. These curves demonstrate that the SI-LSM-AMP algorithm can converge to a good reconstructed result in a reasonable amount of time.

5.4. Visual Quality and Runtime Evaluation in Higher-Power Noise Environment. In this subsection, we conduct similar experiments with noisy CS measurements subject to strong measurement noise to demonstrate the robustness of the proposed SI-LSM-AMP to noise. The standard derivation of additive Gaussian noise is 15. The subjective quality comparison results on Baboon image and Boat image are shown in Figures 11 and 12, respectively. One can observe that (1) the quality of all reconstructed images degrades seriously as measurement noise increases; (2) compared to the reconstructed images in the presence of lower-power measurement noise, the ones in higher-power noise environment tend to be more noisy (e.g., AMP, Turbo-AMP, WaTMRI, and NLR-CS) or smooth (e.g., BM3D-AMP and NL-LSM-AMP); however, the SI-LSM-AMP algorithm can better remove the artifacts and preserve important image structures more effectively even when the measurement noise is high. The corresponding PSNR and SSIM results of these algorithms are also provided, from which we can see that the proposed algorithm achieves the highest quantitative results. This indicates that the SI-LSM-AMP algorithm is shown to be more robust to noise.

The corresponding CPU time versus PSNR curves and iteration number versus PSNR curves are given in Figures 13 and 14. From these figures, we can see that all algorithms converge to reconstructed results more quickly in higher-power noise environment. Among them, the PSNR results of the proposed SI-LSM-AMP algorithm are higher than all other competing methods.

We could find that, compared to other nonlocal methods, the SI-LSM-AMP algorithm has great superiority in higher-power noise environment. These results are reasonable because the mismatching issue is worse when the measurement noise is higher, leading to the difficulty in finding similar patches. However, other nonlocal algorithms depend on self-similarity structure only, which makes them hard to perform well. In contrast, besides the self-similarity structure, the SI-LSM-AMP algorithm also uses the local and global sparse prior of natural images and thus has a soft assumption of the self-similarity structure. As a result, one advantage of the SI-LSM-AMP algorithm is that it is more appropriate for compressive imaging under severe environment, where compressive samples are subject to strong measurement noise.

6. Conclusion

We have proposed an effective structured AMP algorithm for CS image reconstruction. Our work has the following contributions. First, guided by structure sparsity theories, we introduce the Laplacian scale mixture distribution to model the nonlocal sparsity for higher-order sparse representation of natural images, and then it is used as a prior constraint for

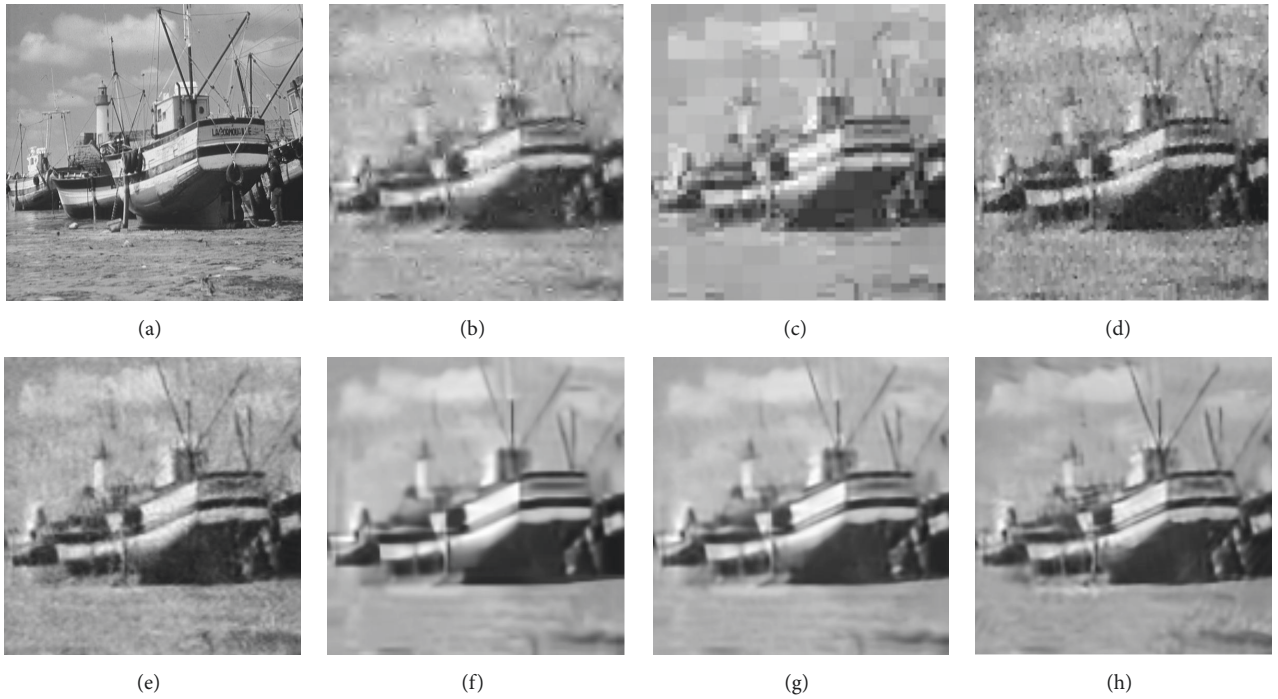


Figure 12: Visual comparisons on Boat image at 20% sampling ratio with measurement noise with standard deviation 15. (a) The original image; (b) AMP (PSNR = 24.34 dB, SSIM = 0.6725); (c) Turbo-AMP (PSNR = 23.24 dB, SSIM = 0.6235); (d) WaTMRI (PSNR = 24.25 dB, SSIM = 0.6793); (e) NLR-CS (PSNR = 26.00 dB, SSIM = 0.7398); (f) BM3D-AMP (PSNR = 26.17 dB, SSIM = 0.7682); (g) NL-LSM-AMP (PSNR = 26.58 dB, SSIM = 0.7895); (h) SI-LSM-AMP (PSNR = 26.84 dB, SSIM = 0.8021).

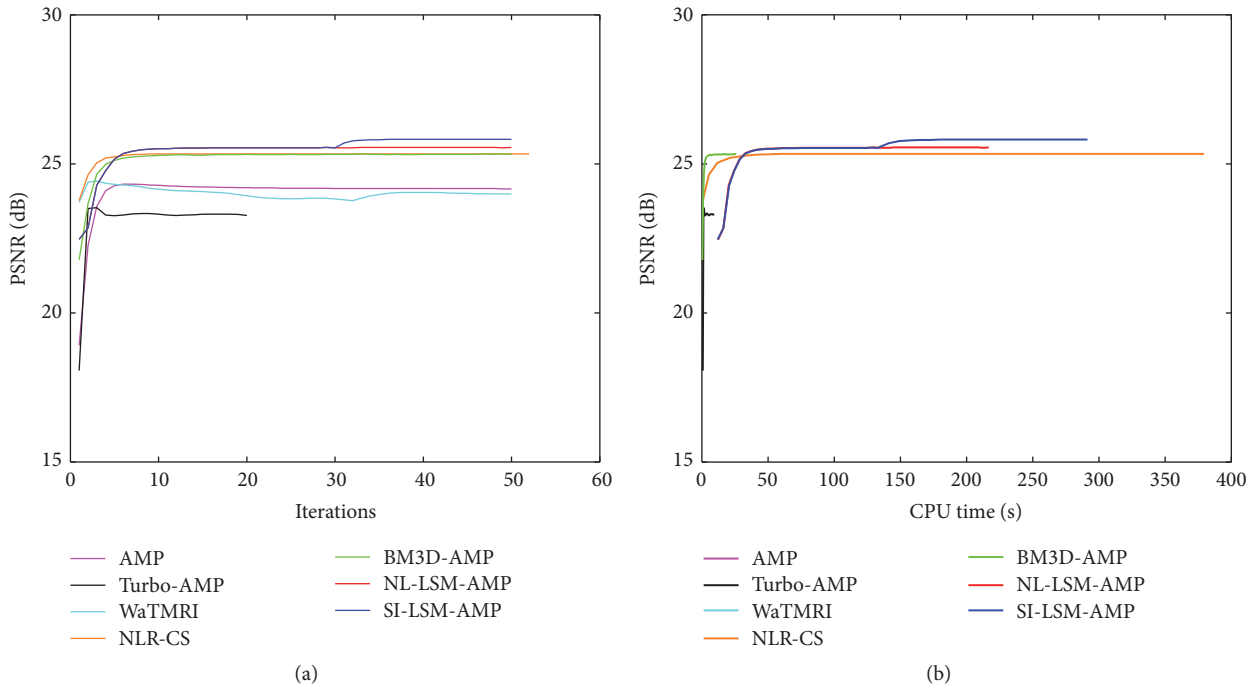


Figure 13: Iterative curves on Baboon image at 20% sampling ratio with measurement noise with standard deviation 15. (a) Average PSNR to iterations; (b) average PSNR to CPU time.

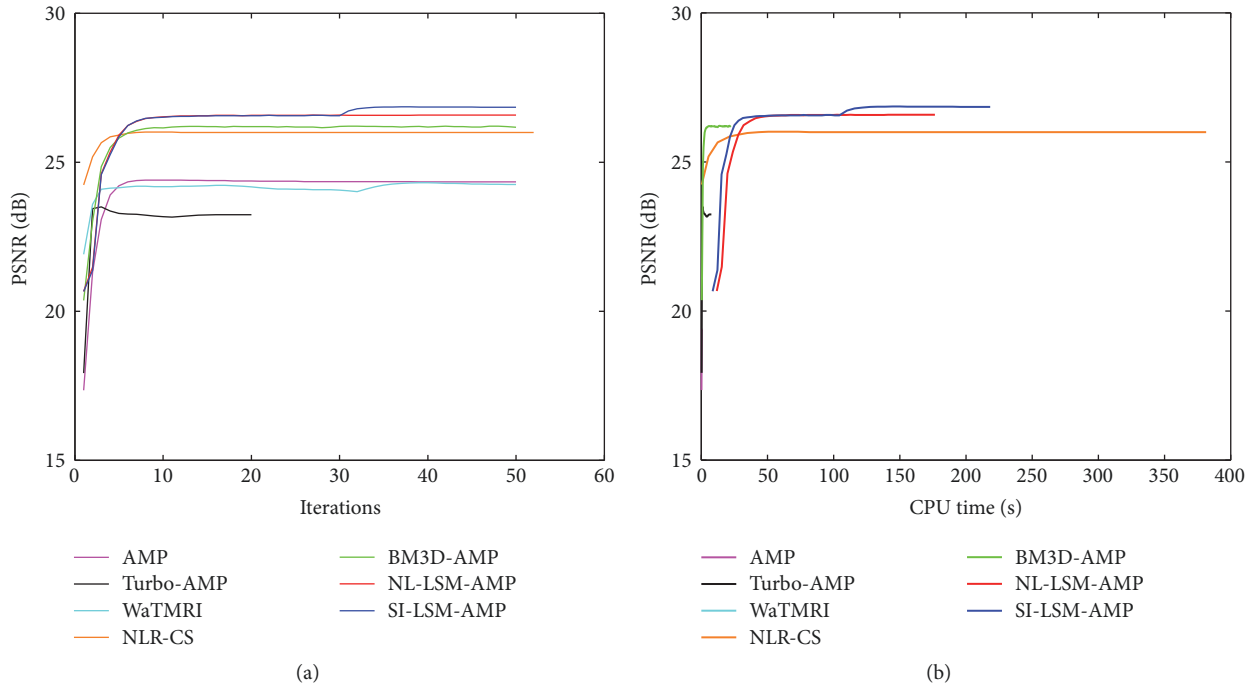


FIGURE 14: Iterative curves on Boat image at 20% sampling ratio with measurement noise with standard deviation 15. (a) Average PSNR to iterations; (b) average PSNR to CPU time.

CS problem. By taking the scale parameter as a random variable, it makes the practical representation much more feasible for achieving a better spatial adaptation. Second, in order to maintain differences between packed patches for irregular structures, we combine local sparsity and global sparsity to soften and complement the nonlocal self-similarity structure assumption. It can substantially enhance the details of image. Afterward, an effective algorithm based on the EM and AMP frame is proposed in this paper to solve this model. The derived inference procedures are efficient to estimate both the sparse coefficients and the scale parameter. After learning the prior parameters, the AMP algorithm is utilized to solve the singular value minimization problem for achieving the accurate image reconstruction. Finally, our simulations and experiments on a variety of natural images demonstrate the superiority of the proposed algorithm to the original AMP algorithm and several tree-based and nonlocal sparsity-based algorithms or solvers in CS image recovery.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was supported by the National Natural Science Foundation of China (no. 61471173), National Engineering Center for Mobile Ultrasonic Detection (no. 2013FU125X02), the Priority Academic Program Development of Jiangsu Higher Education Institutions, and Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology (no. KJR16237).

References

[1] D. L. Donoho, "Compressed sensing," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[3] W. Ke, X. Zhang, Y. Yuan, and J. Shao, "Compressing sensing based source localization for controlled acoustic signals using distributed microphone arrays," *Mathematical Problems in Engineering*, vol. 2017, Article ID 1981280, 11 pages, 2017.

[4] J. Wu, F. Liu, L. C. Jiao, X. Wang, and B. Hou, "Multivariate compressive sensing for image reconstruction in the wavelet domain: using scale mixture models," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3483–3494, 2011.

[5] J. Tan, Y. Ma, and D. Baron, "Compressive imaging via approximate message passing with image denoising," *IEEE Transactions on Signal Processing*, vol. 63, no. 8, pp. 2085–2092, 2015.

[6] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, 1992.

[7] J. Huang, S. Zhang, and D. Metaxas, "Efficient MR image reconstruction for compressed MR imaging," *Medical Image Analysis*, vol. 15, no. 5, pp. 670–679, 2011.

[8] W. Dong, G. Shi, X. Li, L. Zhang, and X. Wu, "Image reconstruction with locally adaptive sparsity and nonlocal robust

regularization," *Signal Processing: Image Communication*, vol. 27, no. 10, pp. 1109–1122, 2012.

[9] W. Chen, I. J. Wassell, and M. R. D. Rodrigues, "Dictionary design for distributed compressive sensing," *IEEE Signal Processing Letters*, vol. 22, no. 1, pp. 95–99, 2015.

[10] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, 2010.

[11] G. Zhang and N. Kingsbury, "Variational Bayesian image restoration with group-sparse modeling of wavelet coefficients," *Digital Signal Processing*, vol. 47, pp. 157–168, 2015.

[12] C. Chen and J. Huang, "Exploiting the wavelet structure in compressed sensing MRI," *Magnetic Resonance Imaging*, vol. 32, no. 10, pp. 1377–1389, 2014.

[13] S. Som and P. Schniter, "Compressive imaging using approximate message passing and a Markov-tree prior," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3439–3448, 2012.

[14] M. Wang, X. Wu, W. Jing, and X. He, "Reconstruction algorithm using exact tree projection for tree-structured compressive sensing," *IET Signal Processing*, vol. 10, no. 5, pp. 566–573, 2016.

[15] X. Zhang, T. Bai, H. Meng, and J. Chen, "Compressive sensing-based ISAR imaging via the combination of the sparsity and nonlocal total variation," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 5, pp. 990–994, 2014.

[16] W. Dong, G. Shi, X. Wu, and L. Zhang, "A learning-based method for compressive image recovery," *Journal of Visual Communication and Image Representation*, vol. 24, no. 7, pp. 1055–1063, 2013.

[17] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618–3632, 2014.

[18] C. A. Metzler, A. Maleki, and R. . Baraniuk, "From denoising to compressed sensing," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 62, no. 9, pp. 5117–5144, 2016.

[19] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "BM3D-AMP: a new image recovery algorithm based on BM3D denoising," in *Proceedings of the IEEE International Conference on Image Processing, ICIP 2015*, pp. 3116–3120, Canada, September 2015.

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

[21] P. Garrigues and B. A. Olshausen, "Group sparse coding with a Laplacian scale mixture prior," in *Proc. Advances in Neural Information Processing Systems*, pp. 676–684, 2010.

[22] A. S. Charles and C. J. Rozell, "Re-weighted l_1 dynamic filtering for time-varying sparse signal estimation," https://arxiv.org/abs/1208.0325.

[23] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–17, 1977.

[24] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 106, no. 45, pp. 18914–18919, 2009.

[25] J. P. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4658–4672, 2013.

[26] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[27] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[28] X. Zhu, "Approximately normalized iterative hard thresholding for nonlinear compressive sensing," *Mathematical Problems in Engineering*, vol. 2016, Article ID 2594752, 2016.

[29] M. S. Hosseini and K. N. Plataniotis, "High-accuracy total variation with application to compressed video sensing," *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 3869–3884, 2014.

[30] Q. Ling, W. Shi, G. Wu, and A. Ribeiro, "DLM: decentralized linearized alternating direction method of multipliers," *IEEE Transactions on Signal Processing*, vol. 63, no. 15, pp. 4051–4064, 2015.

[31] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for l_1-minimization with applications to compressed sensing," *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.

[32] T. Qiao, W. Li, and B. Wu, "A new algorithm based on linearized Bregman iteration with generalized inverse for compressed sensing," *Circuits, Systems and Signal Processing*, vol. 33, no. 5, pp. 1527–1539, 2014.

[33] X. Wang and J. Liang, "Side information-aided compressed sensing reconstruction via approximate message passing," in *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2014*, pp. 3330–3334, Italy, May 2014.

[34] D. L. Donoho, I. Johnstone, and A. Montanari, "Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising," *Institute of Electrical and Electronics Engineers Transactions on Information Theory*, vol. 59, no. 6, pp. 3396–3433, 2013.

[35] P. R. Hill, J.-H. Kim, A. Basarab, D. Kouame, D. R. Bull, and A. Achim, "Compressive imaging using approximate message passing and a Cauchy prior in the wavelet domain," in *Proceedings of the 23rd IEEE International Conference on Image Processing, ICIP 2016*, pp. 2514–2518, USA, September 2016.

[36] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: a low-rank approach," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 700–711, 2013.

[37] Y. Romano and M. Elad, "Boosting of image denoising algorithms," *SIAM Journal on Imaging Sciences*, vol. 8, no. 2, pp. 1187–1219, 2015.

[38] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 2862–2869, June 2014.

[39] A. Montanari, "Graphical models concepts in compressed sensing," https://arxiv.org/abs/1011.4328.

[40] S. Ramani, T. Blu, and M. Unser, "Monte-Carlo sure: a black-box optimization of regularization parameters for general denoising algorithms," *IEEE Transactions on Image Processing*, vol. 17, no. 9, pp. 1540–1554, 2008.