Nonlocal Image Restoration With Bilateral Variance Estimation: A Low-Rank Approach

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

***Abstract*— Simultaneous sparse coding (SSC) or nonlocal image representation has shown great potential in various low-level vision tasks, leading to several state-of-the-art image restoration techniques, including BM3D and LSSC. However, it still lacks a physically plausible explanation about why SSC is a better model than conventional sparse coding for the class of natural images. Meanwhile, the problem of sparsity optimization, especially when tangled with dictionary learning, is computationally difficult to solve. In this paper, we take a low-rank approach toward SSC and provide a conceptually simple interpretation from a bilateral variance estimation perspective, namely that singular-value decomposition of similar packed patches can be viewed as pooling both local and nonlocal information for estimating signal variances. Such perspective inspires us to develop a new class of image restoration algorithms called spatially adaptive iterative singular-value thresholding (SAIST). For noise data, SAIST generalizes the celebrated BayesShrink from local to nonlocal models; for incomplete data, SAIST extends previous deterministic annealing-based solution to sparsity optimization through incorporating the idea of dictionary learning. In addition to conceptual simplicity and computational efficiency, SAIST has achieved highly competent (often better) objective performance compared to several state-of-the-art methods in image denoising and completion experiments. Our subjective quality results compare favorably with those obtained by existing techniques, especially at high noise levels and with a large amount of missing data.**

***Index Terms*— Deterministic annealing, iterative regularization, low-rank method, simultaneous sparse coding, singular-value thresholding.**

I. Introduction

NONLOCAL image representations have received increasingly more attention in recent years. Since the appearance of nonlocal-mean denoising [1], a flurry of advanced nonlocal image restoration algorithms have been developed [2]–[12]. Among them a denoising algorithm named block-matching 3D filtering (BM3D) [3] is likely to be the most well-known due to its outstanding experimental performance. However, the mechanism of nonlocal image denoising remains elusive[1] - e.g., why exploiting nonlocal similarity is a good idea? how to gain a deeper understanding of sparsity under a nonlocal framework? and more importantly, how to further improve the performance of nonlocal image restoration algorithms?

We attempt to partially answer those questions by adapting an old folk-song (low-rank methods) to a fast new tune (nonlocal image restoration) [13], [14]. Despite the long history of low-rank methods (e.g., excellent energy compaction property of SVD has been known for decades [15]), its computational complexity used to be thought of prohibitive [16]. Only recently, low-rank methods were rediscovered "on the heels of compressed sensing" [14] and successfully applied to various matrix completion problems including collaborative filtering [13], image alignment [17], shadow removal [18] and video denoising [9]. However, the connection between nonlocal image models and low-rank methods has largely remained elusive in the open literature to the best of our knowledge. To fill in this gap, our contributions in this paper are two-fold.

On one hand, we propose a powerful image model in the patch space that connects low-rank methods with simultaneous sparse coding. For the first time, we demonstrate an interesting relationship between singular values of a data matrix (likelihood term) and pseudo-metric norm $||A||_{1,2}$ (prior term) in simultaneous sparse coding [7]. Such connection allows us to obtain a novel interpretation of singular value decomposition (SVD) from a *bilateral variance estimation* perspective. It will be shown for a data matrix consisting of similar patches, left-multiplying and right-multiplying matrices of SVD jointly characterize the *local variation* in the row space and *nonlocal variation* in the column space respectively. Under the context of image modeling, we argue that local and nonlocal variations are the two sides of the same coin; therefore it is important to strike a good balance between them.

On the other hand, the proposed image model allows us to cast various image restoration tasks as Bayesian inference problems and solve them in a principled fashion. Our approach unifies previously known methods for noisy (e.g., BayesShrink [19], robust principal component

Manuscript received February 13, 2012; revised September 7, 2012; accepted September 17, 2012. Date of publication October 2, 2012; date of current version January 10, 2013. This work was supported in part by the Natural Science Foundation under Grant CCF-0914353 and Grant ECCS-0968730, the Natural Science Foundation of China under Grant 61100154, Grant 61033004, Grant 61070138, and Grant 61072104, the Major State Basic Research Development Program of China 973 Program under Grant 2013CB329402, and the Fundamental Research Funds of the Central Universities of China under Grant K50510020003. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Debargha Mukherjee.

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

X. Li is with the Lane Department of CSEE, West Virginia University, Morgantown, WV 26506-6109 USA (e-mail: xin.li@ieee.org).

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

Digital Object Identifier 10.1109/TIP.2012.2221729

[1]Private communication with E. Simoncelli.

analysis [20]) or incomplete (e.g., Bayesian inpainting [21], Bayesian compressed sensing [22], [23]) observations and extends them from local to nonlocal models. Moreover, our analysis suggests an intrinsic connection between deterministic annealing [24] and iterative regularization [25] - namely, both of them can be interpreted as the strategies of achieving spatial adaptation. We argue that the strategy of spatial adaptation is critical to the art of tuning the proposed image model to real-world data (i.e., natural images). The newly developed algorithm - dubbed *S*patially *A*daptive *I*terative *S*ingular-value *T*hresholding (SAIST) - has achieved highly competitive (and often better) performance in both tasks of image denoising and image completion.

The rest of this paper is organized as follows. In Sec. II (model), we present a low-rank approach toward modeling nonlocal similarity in natural images and discuss its connection with existing idea of simultaneous sparse coding. In Sec. III (algorithm), we develop our SAIST algorithm borrowing ideas of iterative regularization and deterministic annealing. In Sec. IV (applications), we report our experimental results to demonstrate the excellent performance of SAIST in two scenarios: image denoising and image completion. In Sec. V (connections), we discuss the issues of sparsity in signal modeling and convexity in optimization algorithms. Some concluding remarks are included in Sec. VI.

II. Modeling Nonlocal Similarity of Natural Images: A Low-Rank Approach

A. Simultaneous Sparse Coding via Low-Rank Approximation

Sparse coding in the patch space attempts to represent an image $\boldsymbol{x}$ by a dictionary $\mathbf{U}$[2] and a collection of sparse vectors $\boldsymbol{\alpha}_i$'s - namely

$$\boldsymbol{x} \approx \mathbf{U} \circ \boldsymbol{\alpha} \triangleq \left(\sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{R}_i\right)^{-1} \sum_{i=1}^{N} (\mathbf{R}_i^T \mathbf{U} \boldsymbol{\alpha}_i) \tag{1}$$

where $\mathbf{R}_i$ denotes a matrix extracting an image patch from the i-th position. Various image restoration tasks can be formulated into the following minimization problem [26]

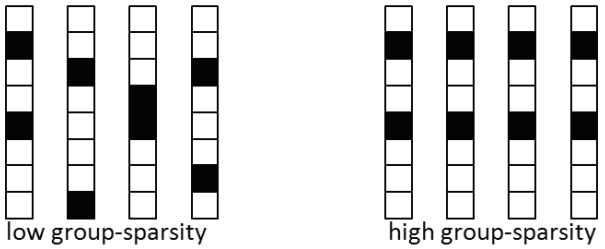
$$(\mathbf{U}, \boldsymbol{\alpha}) = \underset{\mathbf{U},\boldsymbol{\alpha}}{\operatorname{argmin}} \sum_{i=1}^{N} ||\boldsymbol{y}_i - \mathbf{U}\boldsymbol{\alpha}_i||_2^2 + \tau \sum_{i}^{N} ||\boldsymbol{\alpha}_i||_1 \tag{2}$$

where $\boldsymbol{y}_i \in R^n$ (n-patch size) denotes the degraded version of $\boldsymbol{x}_i$, $\boldsymbol{\alpha}_i \in R^n$ is the sparse code with respect to dictionary $\mathbf{U}$ and τ is the regularization parameter.

A major drawback of Eq. (2) is the assumption about the *independence* between sparsely-coded patches. To better exploit the dependencies among patches, the basic idea of *simultaneous* sparse coding [7] - also known as group/structured sparsity [27] - is to group a set of similar patches $\mathbf{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m] \in R^{n \times m}$ (e.g., $m = k + 1$ if one finds the k-nearest-neighbor of an exemplar patch $\boldsymbol{y}_1$) and consider the *group sparsity* defined by a pseudo-matrix norm $||\mathbf{A}||_{p,q}$:

$$(\mathbf{U}, \mathbf{A}) = \underset{\mathbf{A}}{\operatorname{argmin}} ||\mathbf{Y} - \mathbf{U}\mathbf{A}||_F^2 + \tau ||\mathbf{A}||_{p,q} \tag{3}$$

[2]The dictionary can be either fixed such as discrete-cosine-transform or learned through principal component analysis.



(a) (b)

Fig. 1. Art of modeling image evolves from (a) SC to (b) SSC.

where $\mathbf{A} = [\boldsymbol{\alpha}^1; \boldsymbol{\alpha}^2, \ldots, \boldsymbol{\alpha}^n]$ is related to image patches by $\mathbf{X} = \mathbf{U}\mathbf{A}$. We note that the pseudo-matrix norm $||\cdot||_{p,q}$ is defined by [28]

$$||\mathbf{A}||_{p,q} \triangleq \sum_{i=1}^{n} ||\boldsymbol{\alpha}^i||_q^p \tag{4}$$

where $\boldsymbol{\alpha}^i = [\alpha_{i,1}, \ldots, \alpha_{i,m}]$ denotes the i^{th} row of matrix $\mathbf{A}$ in $R^{n \times m}$. However, unlike the original sparsity (please refer to Fig. 1), we note that the above formulation of group-sparsity encourages the *alignment* of sparse coefficients along the row direction only. In other words, it does not treat the row and column spaces equally - namely, a matrix $\mathbf{A}$ and its transpose $\mathbf{A}^T$ will be characterized by varying amount of group-sparsity (though they have the same amount of sparsity). This is undesirable because the row and column spaces of $\mathbf{A}$ respectively characterize the nonlocal and local variations associated with the exemplar patch. There is no prior knowledge to favor either local or nonlocal view during the formulation of group-sparsity.

One possible approach of restoring the symmetry between row and column spaces is to introduce a right-multiplying matrix $\mathbf{V}$ and rewrite the matrix $\mathbf{A}$ into

$$\mathbf{A} = \Sigma \mathbf{V}^T \tag{5}$$

where $\Sigma = diag\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ ($K = min(m, n)$) is a diagonal matrix in $R^{K \times K}$ and each column of $\mathbf{V}$ in $R^{m \times K}$ is decomposed of $\boldsymbol{v}_i = \frac{1}{\lambda_i}(\boldsymbol{\alpha}^i)^T$. In other words, the new sparsifying matrix $\mathbf{V}$ plays the dual role of the dictionary $\mathbf{U}$ (one works with rows and the other with columns). Therefore, Eq. (3) can be rewritten into

$$(\mathbf{U}, \Sigma, \mathbf{V}) = \underset{\mathbf{U},\Sigma,\mathbf{V}}{\operatorname{argmin}} ||\mathbf{Y} - \mathbf{U}\Sigma\mathbf{V}^T||_F^2 + \tau ||\mathbf{A}||_{p,q}. \tag{6}$$

We note that the connection between data term and regularization term is not obvious at this point; it remains to interpret $||\mathbf{A}||_{p,q}$ in the newly-defined Σ-domain as we will elaborate next.

The key observation behind our approach is that when $p = 1, q = 2$, the group sparsity regularizer $||\mathbf{A}||_{1,2}$ in fact computes the sum of standard deviations associated with sparse coefficient vector in each row, i.e.,

$$||\mathbf{A}||_{1,2} = \sum_{i=1}^{K} \sqrt{\alpha_{i,1}^2 + \alpha_{i,2}^2 + \cdots + \alpha_{i,m}^2} = \sum_{i=1}^{K} \sqrt{m}\sigma_i \tag{7}$$

where σ_i denotes the standard deviation of the sparse coefficients $\boldsymbol{\alpha}^i$ in the i-th row. Therefore, we have $\boldsymbol{\alpha}^i = \lambda_i \boldsymbol{v}_i^T$

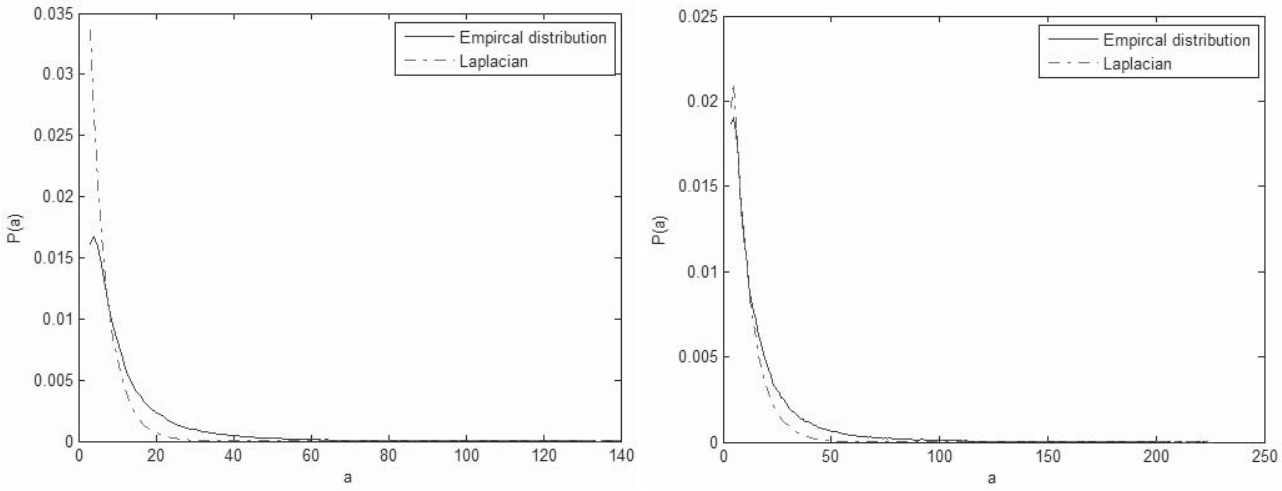


Fig. 2. Fitting the empirical distribution of σ_i (standard deviation of sparse coefficients $\boldsymbol{\alpha}^i$) by Laplacian on two standard test images. (a) *Lena*. (b) *Monarch*.

and

$$\sigma_i^2 = \frac{1}{m}||\boldsymbol{\alpha}^i||_2^2 = \frac{1}{m}||\lambda_i \boldsymbol{v}_i^T||_2^2 = \frac{\lambda_i^2}{m} \tag{8}$$

where the first identity comes from Eq. (7) and the last one is due to the unitary property of $\mathbf{V}$. The significance of Eq. (8) lies in its implication that any operation designed for sparse coefficient vector $\boldsymbol{\alpha}$'s can be conveniently implemented with σ_i or λ_i (only differs by a constant scalar). Substituting Eq. (8) into Eq. (6), we obtain

$$(\mathbf{U}, \Sigma, \mathbf{V}) = \underset{\mathbf{U},\Sigma,\mathbf{V}}{\operatorname{argmin}} ||\mathbf{Y} - \mathbf{U}\Sigma\mathbf{V}^T||_F^2 + \tau \sum_{i=1}^{K} \lambda_i \tag{9}$$

which is a standard low-rank approximation problem [13] and admits the following singular value decomposition (SVD)-based solution [13]

$$\begin{cases} (\mathbf{U}, \Sigma, \mathbf{V}) = \operatorname{svd}(\mathbf{Y}); \\ \hat{\Sigma} = \boldsymbol{S}_\tau(\Sigma); \end{cases} \tag{10}$$

where $\boldsymbol{S}_\tau$ denotes the soft thresholding operator with threshold τ (regularization parameter) and the reconstructed data matrix is conveniently obtained by $\hat{\mathbf{X}} = \mathbf{U}\hat{\Sigma}\mathbf{V}^T$.

B. Singular Value Decomposition as Bilateral Variance Estimation

To gain a deeper understanding of the above SVD-based approach, we propose to study the role of left-multiplying and right-multiplying matrices $(\mathbf{U}, \mathbf{V})$ from a signal variance estimation perspective. An image is viewed as a mapping from spatial domain $[1, H] \times [1, W]$ to intensity range $[0, 255]$. In local image models (e.g., wavelet-based or DCT-based), signal variance estimation is often based on the *domain*-Markovian assumption (e.g., [29]) - namely, the probability distribution function of a coefficient - when conditioned on a local window (its spatial neighborhood) - is independent of the coefficients outside the window. The legitimacy of such assumption is largely dependent on the *transient* behavior of the signal as well as the locality property of basis functions [30]. For example, a basis function with short support is often preferred for sharp edges; while slowly-varying signals call for a basis function with long support.

Alternatively, one can make a *range*-Markovian assumption about the image source as implicitly made in a flurry of nonlocal image denoising algorithms (e.g., [1], [3], [11]). More specifically, the local window does not refer to the spatial neighborhood of a pixel/coefficient but the k-Nearest-Neighbor (kNN) of the image patch associated with that pixel[3]. It follows that the procedure of variance estimation (equivalently the calculation of standard deviation from a local window) can also be done nonlocally with respect to kNN rather than locally within a spatial neighborhood. We claim that such idea of *nonlocal variance estimation* is implemented by the right-multiplying matrix $\mathbf{V}$ in our low-rank approach.

Let us recall $\mathbf{X} = \mathbf{U}\mathbf{A}$ where $\mathbf{U}$ is the dictionary of orthogonal local-basis functions and $\mathbf{A} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m]$ is the sparse coefficients associated with the kNN of an exemplar patch. Also note that in our notation subscripts and superscripts respectively denote the column and row vectors of matrix $\mathbf{A}$. If the sparse coefficient vector of each row $\boldsymbol{\alpha}^i$ is assumed to observe a Gaussian distribution $N(0, \sigma_i^2)$, how would one calculate its standard deviation? A common approach is as follows

$$\hat{\sigma}_i = \frac{1}{m}\sqrt{\sum_{j=1}^{m} \alpha_{i,j}^2}. \tag{11}$$

Comparing Eq. (11) with Eqs. (7) and (8), we can see that the singular values of Σ can be interpreted as the result of nonlocal variance estimation. In other words, the diagonalization of $\mathbf{A}$ can be viewed as a *parallel* implementation of calculating the standard deviations for sparse coefficient vectors of n rows thanks to the unitary property of right-multiplying matrix $\mathbf{V}$. To summarize, left-multiplying and right-multiplying matrices

[3]Indeed any pixel could be associated with multiple patches; such issue is related to the redundancy of patch-based image models and has been addressed under the framework of Bayesian model averaging in [31].

Algorithm 1 Image Denoising via SAIST

- Initialization: $\hat{\boldsymbol{x}} = \boldsymbol{y}$;
- Iterate on $i = 1, 2, \ldots, iter$
  - Patch clustering: find the kNN for each exemplar patch and create data matrices $\mathbf{Y}_i$'s for each cluster;
  - Iterative regularization: $\boldsymbol{y}^{(k+1)} = \hat{\boldsymbol{x}}^{(k)} + \delta(\boldsymbol{y} - \hat{\boldsymbol{x}}^{(k)})$;
  - Noise variance update: re-estimate σ_w^2 from $\boldsymbol{y}^{(k+1)}$ via Eq. 15;
  - SVD for each noisy data matrix $\mathbf{Y}_i$: $(\mathbf{U}_i, \Sigma_i, \mathbf{V}_i) = \text{svd}(\mathbf{Y}_i)$;
  - Thresholds update: compute τ_i using Eqs. (12) and (13);
  - Singular value thresholding: $\hat{\lambda}_i = \boldsymbol{S}_{\tau_i}(\lambda_i)$ with computed τ_i;
  - Image update: obtain an improved denoised image $\hat{\boldsymbol{x}}^{(k)}$ by weighted averaging all denoised patches $\hat{\mathbf{X}}_i = \mathbf{U}_i \hat{\Sigma}_i \mathbf{V}_i^T$;
- Output: $\hat{\boldsymbol{x}}^{(k)}$.

$(\mathbf{U}, \mathbf{V})$ respectively characterize the local (column-wise) and nonlocal (row-wise) variations of data matrix $\mathbf{X}$. We argue that such a bilateral interpretation of SVD is physically more plausible and important to our conceptual understanding of the match between the proposed image model and real-world data.

How would the empirical distribution of singular values λ_i's look like? Fig. 2 shows the fitting result of empirical distributions of $\boldsymbol{\sigma}$ on two typical test images: *lena* and *monarch*. It can be observed that both empirical distributions can be reasonably well approximated by Laplacian - we note that such statistical modeling experiment is analogous to well-known studies on wavelet coefficients (e.g., [32]) but has to be understood under a different context (simultaneous sparse coding instead of sparse coding) because each data matrix $\mathbf{X}$ is decomposed of *nonlocal* similar patches - it shows the prevalence of heavy-tail distribution [33] in nonlocal sparse representations too.

III. Iterative Singular-Value Thresholding With Deterministic Annealing

In this section, we develop a class of image restoration algorithms based on nonlocal sparse representations discussed above. Under the framework of Bayesian inference, we focus on two types of observation models (or likelihood functions): *noisy* and *incomplete* data. As the observation model changes, the strategy of choosing the threshold varies but share a common objective of achieving spatial adaption. As we will see, it is enlightening to understand recently proposed iterative regularization techniques from the spatially adaptive perspective - namely, different classes of image structures correspond to different types of saddle points and iterative regularization represents a deterministic annealing (DA)-like strategy of traversing the image space.

A. Image Restoration from Noisy Observation Data

Given noisy observation $\mathbf{Y} = \mathbf{X} + \mathbf{W}$, we note that studies on wavelet-based image denoising (e.g., [19]) have shown the following thumb rule for choosing the threshold: $\tau = 2\sqrt{2}\sigma_w^2/\sigma_x$ under the assumption that wavelet coefficients observe i.i.d. Laplacian distribution. As shown in Fig. 2, the empirical distribution of singular values can also be modeled by a Laplacian with spatially changing variances. With a spatially adaptive Laplacian prior, we can set the threshold parameter to be [19]

$$\tau_i = \frac{2\sqrt{2}\sigma_w^2}{\sigma_i} \tag{12}$$

where σ_i denotes the locally estimated variance at the position i. Following the arguments of bilateral variance estimation, we can show that the one-sample maximum-likelihood (ML) estimation of signal variance σ_i is given by [19]

$$\hat{\sigma}_i = \sqrt{max(\tilde{\lambda}_i^2/m - \sigma_w^2, 0)} \tag{13}$$

where $\tilde{\lambda}_i$ denotes the singular-value calculated from the noisy data matrix $\mathbf{Y}$. In fact, the above formula can be interpreted as the bilateral extension of BayesShrink [19] from wavelet-domain (local information only) to SVD-domain (exploiting both local and nonlocal information).

Additionally, we suggest that recently developed iterative regularization techniques [25] offers an alternative approach toward spatial adaptation. The basic idea of iterative regularization is to add filtered noise back to the denoised image - i.e.,

$$\boldsymbol{y}^{(k+1)} = \hat{\boldsymbol{x}}^{(k)} + \delta(\boldsymbol{y} - \hat{\boldsymbol{x}}^{(k)}) \tag{14}$$

where k denotes the iteration number and δ is a relaxation parameter. Performance improvement by iterative regularization has been reported for wavelet-based [34], total-variation and bilateral filtering models [35]. Here, we propose to extend the idea of iterative regularization to update the estimation of noise variance (due to the feedback of filtered noise) and signal variance (due to the update of clustering results) alternatively - i.e.,

$$\hat{\sigma}_w^{(k)} = \gamma \sqrt{\sigma_w^2 - ||\boldsymbol{y} - \boldsymbol{y}^{(k+1)}||_{l_2}} \tag{15}$$

where γ is a scaling factor controlling the re-estimation of noise variance and

$$\hat{\sigma}_i^{(k+1)} = \sqrt{max((\tilde{\lambda}_i^{(k)})^2/m - (\hat{\sigma}_w^{(k)})^2, 0)}. \tag{16}$$

An intuitive explanation of Eqs. (15) and (16) is as follows. As the iteration starts, only strong signals (with large singular values) can survive soft-thresholding and contribute to the initial guess of $\hat{\boldsymbol{x}}$; however, partially recovered signal will be fed back to the noisy observation through Eq. (14), which helps lower the estimation of noise. In return, weaker signals can be identified and added to the signal estimate. As the iteration progresses, we usually observe that the estimated noise variance monotonically decreases; meanwhile image structures are progressively recovered until the convergence. To facilitate the conceptual understanding, an animation recording the intermediate denoised image results can be accessed at http://www.csee.wvu.edu/~xinl/demo/saist.html. A step-by-step description of SAIST denoising algorithm is given below.

The implementation of SAIST denoising algorithm (Algorithm 1) under MATLAB is relatively easy - only two parameters need to be specified by the user: the patch size and the

Algorithm 2 Image Completion via SAIST

- Initialization: set initial estimate $\hat{\boldsymbol{x}}^{(0)}$ and threshold $\tau = \tau_0$;
- Iterate on $i = 0, 1, 2, \ldots, iter$
  - Patch clustering: find the kNN for each exemplar patch and create data matrices $\mathbf{X}_i$'s for each cluster;
  - Landweber iteration (Φ, Φ^t denote measurement operator and its adjoint): run $\hat{\boldsymbol{x}}^{(k+1)} = \hat{\boldsymbol{x}}^{(k)} + \lambda\Phi^t(\boldsymbol{y} - \Phi\hat{\boldsymbol{x}}^{(k)})$ for r-times;
  - Perform SVD for each data matrix $\mathbf{X}_i$: $(\mathbf{U}_i, \Sigma_i, \mathbf{V}_i) = svd(\mathbf{X}_i)$;
  - Singular value thresholding: $\hat{\lambda}_i = \boldsymbol{S}_{\tau_i}(\lambda_i)$ with computed τ_i;
  - Image update: obtain a new reconstructed image $\hat{\boldsymbol{x}}^{(k)}$ by weighted averaging all reconstructed patches $\hat{\mathbf{X}}_{i+1} = \mathbf{U}_i\hat{\Sigma}_i\mathbf{V}_i^T$;
  - Deterministic annealing: $\tau = (1-\epsilon)^k\tau_0, \forall \mathrm{mod}(k, T) = 0$;
- Output: $\hat{\boldsymbol{x}}^{(k)}$.

number of similar patches. The core component of singular-value thresholding only takes about 10 lines of MATLAB codes. Similar to BM3D [3], we only use exemplars every five pixels along both horizontal and vertical directions to speed up the computation; the final output of each pixel is weighted average across multiple patches containing that pixel. The weight is empirically set as $\omega_j = 1 - s_j/n$, where s_j denotes the number of nonzero singular values of the denoised $\hat{\mathbf{X}}_i$. In our current MATLAB implementation, we have exploited the vectorized functions (e.g., $repmat$) to obtain the kNN search results fast. It takes SAIST less than 20 seconds to denoise a 256×256 image on a Intel Dual Core i7 CPU 2.67 GHz.

B. Image Completion from Incomplete Observation Data

Given incomplete data $\mathbf{Y} = \Phi\mathbf{X}|_\Omega$ (Φ, Ω denote the operator and the support of measurement process), we can observe the interesting duality between matrix completion [13], [36] and image completion (a.k.a. image inpainting [37] and image recovery [38]). However, unlike matrix completion, image completion often involves the filling of missing data in a number of data matrices each corresponding to the kNN of a chosen exemplar patch. Moreover, there is no universal low-rank prior for different class of image structures (e.g., smooth areas vs. edges). Therefore, the technique of singular-value thresholding has to be used with caution when applied to the task of image completion - for example, group-sparsity optimization is often tangled with dictionary learning, which makes the problem nonconvex. Accordingly, several researchers have opted to decouple the two and learn the dictionary from a large set of training data (e.g., [7], [12]). Here, we propose to borrow the strategy of deterministic annealing to tackle such nonconvex sparsity optimization problem.

The idea of deterministic annealing (DA) is simple - one starts with a large threshold and then progressively decrease the threshold value according to some annealing schedule [39]. It was first discovered that DA can dramatically improve the performance of DCT-based [40] and wavelet-based [41] image recovery algorithms as well as BM3D-based compressed sensing [42] and interpolation [43]. Later, the idea of DA was used jointly with hybrid sparse representations [44] and explained from the viewpoint of spatial adaptation. Since the left-multiplying and right-multiplying matrices ($\mathbf{U}$, $\mathbf{V}$) of SVD characterize the local and nonlocal variations of data matrix respectively, it is plausible to extend the idea of DA into singular-value thresholding (SVT). We also give a complete description of SAIST completing algorithm as follows.

It is worth noting that Algorithm 2 is general enough to be applied into different scenarios. When Φ is spatial sampling, Algorithm 2 becomes image inpainting [37] (if Ω^c is set to be the so-called inpainting domain), regularly sampled image interpolation [45] (if Ω is set to be odd-indexed rows and columns) or irregularly sampled image interpolation [46] (if Ω is set to be random samples in the spatial domain). When Φ is Fourier sampling and Ω denotes a collection of radial lines in the Fourier space, Algorithm 2 becomes the image reconstruction tool useful for compressed-sensing [47]. We acknowledge that the parameter setting of Algorithm 2 has to vary from application to application; but as we will show next, the general principle is the same and Algorithm 2 is capable of advancing the state-of-the-art in several applications.

IV. Reproducible Experimental Results

In this section, we report our experimental results with SAIST denoising/completion algorithms described in the previous section. These experimental results are used to support the effectiveness of the proposed image model and the idea of bilateral variance estimation. All the benchmark schemes used in our experimental comparison are the latest works published with reproducible results. To promote the culture of reproducible research, source codes and saved experimental results accompanying this work can be accessed at http://www.csee.wvu.edu/~xinl/demo/saist.html.

A. Image Denoising

We first compare the proposed SAIST denoising algorithm and three leading methods for removing additive white Gaussian noise: block-matching 3D (BM3D) [3], learning simultaneous sparse coding (LSSC) [7] and clustering-based sparse representation (CSR) [10]. It should be noted that LSSC and CSR are the only two schemes in the open literature whose denoising performance has shown convincing improvements over BM3D (published over five years ago). The denoising results of all benchmark schemes are generated from the source codes or executables released by their authors. The PSNR performance of four competing denoising algorithms are reported in Table 1 (the highest PSNR value is highlighted in each cell to facilitate the comparison). From Table 1 we can see that SAIST achieves at least comparable denoising performance to other three methods. On the average, our SAIST outperforms all other three benchmark methods at all noise levels and the gain becomes more significant as the noise level increases.

The visual quality improvements achieved by SAIST seem even more convincing. In Figs. 3–4, we have compared the

TABLE I

COMPARISON OF PSNR (dB) RESULTS AMONG FOUR COMPETING DENOISING METHODS. IN EACH CELL, FOUR DENOISING RESULTS ARE REPORTED. TOP LEFT: BM3D [3]. TOP RIGHT: LSSC [7]. BOTTOM LEFT: CSR [10]. BOTTOM RIGHT: LASSC (THIS PAPER). THE BEST RESULT AMONG FOUR IS HIGHLIGHTED IN EACH CELL

σ	10		15		20		30		50		100	
Lena	**35.94**	35.85	**34.29**	34.16	33.07	32.89	**31.28**	31.19	**29.08**	28.95	25.91	25.96
	35.90	35.90	34.20	34.21	32.96	**33.08**	31.16	31.27	28.79	29.01	25.33	**25.93**
Monarch	34.14	34.49	31.88	32.16	30.38	30.59	28.36	28.08	25.69	25.59	22.33	21.82
	34.49	**34.80**	32.25	**32.52**	30.71	**30.81**	28.56	**28.68**	25.79	**25.89**	21.75	**22.41**
Barbara	34.97	34.97	33.09	32.98	31.74	31.54	29.77	29.62	27.28	27.13	23.62	23.56
	35.10	**35.23**	33.17	**33.32**	31.78	**32.10**	29.72	**30.09**	26.95	**27.54**	23.05	**24.13**
Boat	33.92	**34.01**	32.15	**32.18**	**30.89**	30.87	**29.11**	29.02	**26.84**	26.76	**24.04**	23.94
	33.88	33.91	32.05	32.09	30.78	30.81	28.94	28.93	26.67	26.66	23.58	23.90
C. Man	34.13	34.16	31.91	31.97	30.51	**30.54**	**28.70**	28.64	26.28	**26.36**	23.18	23.14
	34.06	**34.24**	31.89	**32.01**	30.49	30.50	28.64	28.46	26.27	26.19	22.61	**23.22**
Couple	**34.02**	33.97	**32.10**	32.07	**30.75**	30.70	**28.84**	28.71	**26.48**	26.32	**23.56**	23.34
	33.95	33.90	32.00	31.94	30.60	30.61	28.62	28.68	26.20	26.30	23.20	23.35
F. Print	32.53	32.58	30.35	30.32	28.87	28.78	26.88	26.67	**24.59**	24.21	21.58	21.18
	32.70	**32.71**	**30.47**	30.45	28.97	**29.03**	26.95	**26.98**	24.53	24.55	21.09	**21.59**
Hill	33.62	**33.69**	31.88	**31.90**	**30.73**	30.71	**29.14**	29.05	**27.16**	26.99	**24.49**	24.31
	33.66	33.67	31.87	31.85	30.65	30.68	28.97	28.99	26.90	26.94	24.14	24.27
House	36.82	**37.07**	35.07	**35.34**	33.92	**34.16**	32.21	**32.46**	29.73	29.90	26.03	25.63
	36.88	36.67	35.11	34.91	33.86	33.90	32.11	32.39	29.39	**30.20**	25.37	**26.75**
Man	33.94	34.04	31.88	31.99	30.54	**30.61**	**28.81**	28.77	**26.81**	26.73	**24.16**	24.00
	33.96	**34.08**	31.91	**31.99**	30.56	30.60	28.75	28.74	26.68	26.66	23.87	24.02
Peppers	34.72	34.82	32.75	**32.88**	31.31	**31.48**	29.31	**29.38**	26.69	**26.87**	23.23	23.14
	34.64	**34.82**	32.69	32.87	31.25	31.39	29.22	29.33	26.49	26.76	22.34	**23.32**
Straw	30.99	31.39	28.67	28.96	27.10	27.36	24.99	25.19	22.50	22.67	**19.70**	19.50
	31.51	**31.62**	29.14	**29.23**	27.50	**27.61**	25.16	**25.46**	22.52	**22.85**	19.46	19.54
Average	34.15	34.25	32.17	32.24	30.82	30.85	28.95	28.90	26.59	26.54	23.49	23.29
	34.23	**34.30**	32.23	**32.28**	30.84	**30.93**	28.90	**29.00**	26.43	**26.63**	22.98	**23.54**

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

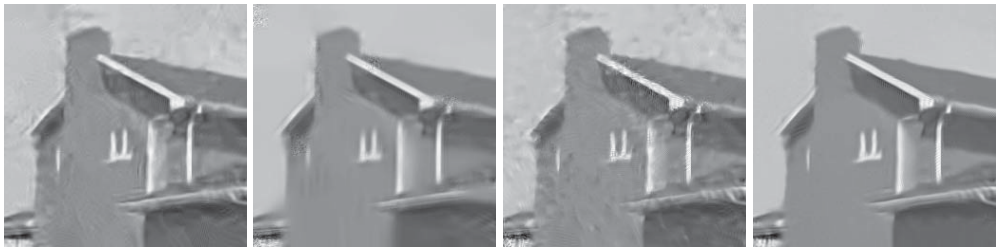
Fig. 3. Denoising performance comparison for *Monarch* image at noise level $\sigma_n = 30$. (a) BM3D [3] (PSNR = 28.36 dB, SSIM = 0.8863). (b) LSSC [7] (PSNR = 28.08 dB, SSIM = 0.8777). (c) CSR [10] (PSNR = 28.56 dB, SSIM = 0.8868). (d) SAIST (PSNR = **28.68** dB, SSIM = **0.8930**).

cropped portions of denoised images by four different methods at two medium noise levels: 30 and 50 - the improvement is noticeable but not striking. Then as we increase the noise level to 100, visual quality difference becomes striking. As can be seen from Figs. 5–6, our SAIST convincingly outperform other denoising methods in the presence of heavy noise. We argue that such advantage can be better appreciated from the signal-variance estimation point of view. According to the Wiener filtering formula, $MMSE \rightarrow \sigma_x^2$ as $\sigma_w \rightarrow \infty$. Therefore, the performance gain achieved by SAIST can be interpreted as a more robust solution to signal-variance estimation.

We have also compared SAIST and BM3D denoising on some more challenging real-world data set. In the real-world scenario, noise associated with raw image data is not additive white Gaussian - e.g., the class of speckle noise is often associated with ultrasound and short-wavelength infrared (SWIR) imaging. Fig. 7 includes the comparison of a phantom image acquired by a SWIR camera and its two

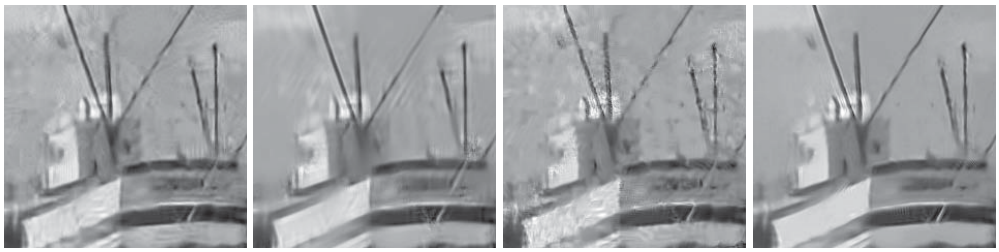
(a) (b) (c) (d)

Fig. 4. Denoising performance comparison for *Barbara* image at noise level $\sigma_n = 50$. (a) BM3D [3] (PSNR = 27.28 dB, SSIM = 0.7964). (b) LSSC [7] (PSNR = 27.13 dB, SSIM = 0.7919). (c) CSR [10] (PSNR = 26.95 dB, SSIM = 0.7895). (d) SAIST (PSNR = **27.54** dB, SSIM = **0.8076**).



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

Fig. 5. Denoising performance comparison for *House* image at noise level $\sigma_n = 100$. (a) BM3D [3] (PSNR = 26.03 dB, SSIM = 0.7146). (b) LSSC [7] (PSNR = 25.63 dB, SSIM = 0.7394). (c) CSR [10] (PSNR = 25.37 dB, SSIM = 0.6866). (d) SAIST (PSNR = **26.75** dB, SSIM = **0.7600**).



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

Fig. 6. Denoising performance comparison for *Boat* image at noise level $\sigma_n = 100$. (a) BM3D [3] (PSNR = **24.04** dB, SSIM = 0.5872). (b) LSSC [7] (PSNR = 23.94 dB, SSIM = 0.5918) (c) CSR [10] (PSNR = 23.58 dB, SSIM = 0.5679). (d) SAIST (PSNR = 23.90 dB, SSIM = **0.5922**).

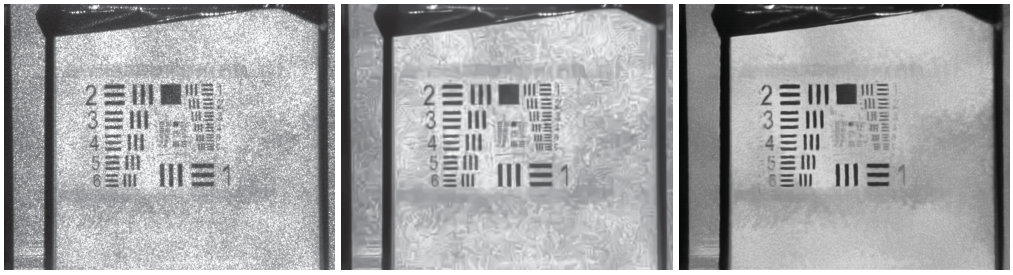
denoised versions (BM3D vs. SAIST)[4]. It can be observed that SAIST is capable of delivering visually more pleasant images from the noisy data than BM3D - in fact, it has been observed before that the performance of BM3D degrades noticeably even in the situation of Gaussian noise when the noise power is high. For example, BM3D has the tendency of being fooled by faulty clustering results in the presence of heavy noise and producing undesirable artifacts in smooth regions. By contrast, the proposed SAIST denoising is much more robust to noise type and strength which seems to make it more appealing in real-world applications.

B. Image Completion

We first report our experimental results for a collection of six small-size (64×64) toy-example images as the representatives of regular edge and texture structures. Our reason for including such comparison is that they facilitate our understanding of the SAIST image completion algorithm from an image modeling perspective. When image size is kept small, it becomes easier to assess the match or mismatch between the model and the data; by contrast, large-size test images are often decomposed of the mixture of different classes of structures. Our experimental setup is identical to that in [40] and [44] - the inpainting domain is the cental 16×16 block and the PSNR is calculated for those missing pixels only. As shown in Fig. 8, we can see that SAIST dramatically outperforms other competing schemes including exemplar-based [48], morphological component analysis (MCA)-based [49] and our own recent work (DA-based [44]. Especially when compared with [44], the key difference lies in the adoption of SVD rather than 2D-FFT as the sparsifying tool.

We have also tested the performance of SAIST-based interpolation in two contrasting settings: irregularly sampled data set and regularly-decimated data set. The experimental setup for image interpolation from irregularly samples is adopted from [46] - we randomly delete 85% pixels from an image and compare the reconstructed images from the remaining 15% samples. It can be observed from Fig. 9 that SAIST can outperform Delaunay-spline interpolation and Iterative steering kernel regression (ISKR) by about 2.2 dB and 1.2 dB

[4]Manual tuning of denoising parameters is necessary for both schemes since noise is not AWGN any more.



(a) (b) (c)

Fig. 7. Denoising performance comparison on real-world data. a) Noisy raw data acquired by a SWIR camera. b) Denoised image given by BM3D. c) Denoised image given by SAIST. Please note the excellent noise suppression result of SAIST in smooth regions.

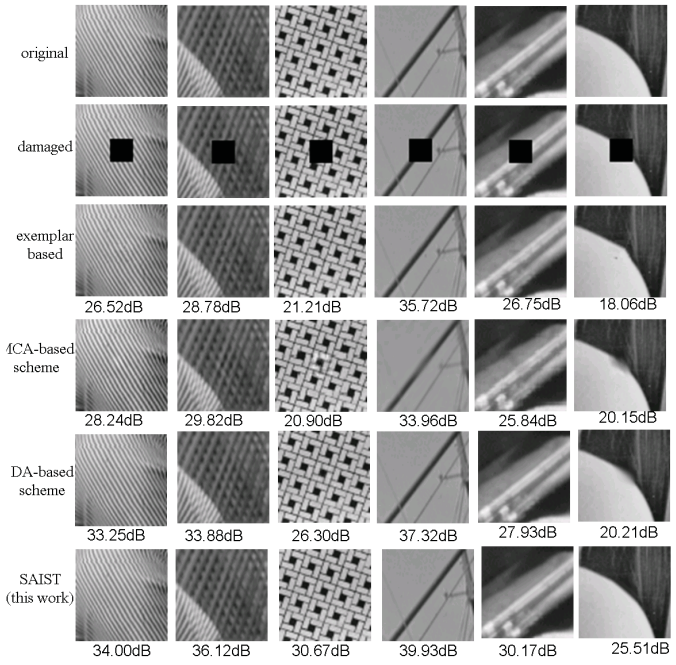


Fig. 8. Inpainting performance comparison among several competing methods: exemplar-based [48], MCA-based [49], DA-based [44], and SAIST-based for six toy-example images (three regular textures and three regular edges).

respectively. Visual quality improvements are also impressive in particular around hair and hat regions. Since ISKR is based on a local model, we conclude that the gain is largely attributed to the exploitation of nonlocal similarity in the given image.

Next, we compare SAIST-based image interpolation and three other leading methods (bicubic, NEDI [45] and SAI [50]) for regularly-decimated samples. It should be noted that 1) all three benchmark methods are based on local models - NEDI improves upon on bicubic by spatially adapting the interpolation coefficients and SAI further improves this idea by enforcing local consistency of autoregressive models; 2) subjective quality of interpolated image is often not faithfully reflected by the PSNR value (e.g., NEDI does not necessarily produces higher PSNR results than bicubic). Referring to Fig. 10, we observe that 1) SAIST can achieve higher PSNR performance than bicubic and NEDI; but falls behind SAI by less than 0.5dB; 2) SAIST seems to be the only one capable of recovering the fine-detail structures in the hat region. To facilitate the visual inspection, we have included the zoomed version of hat region in Fig. 11. It is interesting to see that aliasing artifacts are effectively suppressed in SAIST-interpolated images thanks to the presence of weak yet abundant nonlocal texture patterns (more discussions on this matter can be found in the next section).

Finally, we have compared SAIST and BM3D on a popular experimental setting related to compressed sensing. The challenge is to reconstruct an image from its incomplete samples in Fourier domain. In our experiment, we have adopted the *csphantom* image - an improved version containing more fine-detailed structures than the one created by the matlab function *phantom*. For this image, it is easy to verify that local regularization techniques (e.g., l1-magic) does not work effectively because TV-based model does not represent a good fit with those fine-detailed structures. Fig. 12 includes the PSNR performance comparison between BM3D (to the best of our knowledge, BM3D-CS has achieved the best experimental result of image completion from partial Fourier samples among all reproducible CS software.) and SAIST on *csphantom* image. It can be observed that as much as 11–12 dB is achieved by SAIST over BM3D at the convergence. Due to space limitations, more experimental comparison between SAIST and BM3D can be found at the above-mentioned website accompanying this work.

V. Discussion

A. Modeling: Local Variation Versus Nonlocal Invariance

Is the world transient or invariant? Wavelet theory advocates for the importance of modeling transient events [51] because they carry important information. Recent advances in nonlocal image processing and low-rank methods seem to suggest otherwise - one has to model invariant events in order to gain a deeper understanding of transient ones. Mathematical formulation of changes by local derivatives and memory by Markovian assumptions are powerful tools but should be used with caution. Classical physics might be based on the idea that nature can be described locally [52]; but as the complexity of physical systems increases such that long-range interactions prevail in chemical and biological systems, nonlocal view becomes necessary (e.g., reaction-diffusion systems [53], [54]). Since photographic images are digital representations of natural scenes, they inherit those nonlocal invariant properties including self-similarity [55] and scale-invariance [56].

How do we unify the local and nonlocal views in a common framework? We argue that low-rank approximation

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

Fig. 9. Image reconstruction performance comparison from irregular samples for *Lena* image. (a) Irregularly sampled data set (15% of pixels at random locations are preserved). (b) Delaunay-spline interpolation (PSNR = 28.86 dB, SSIM = 0.8419). (c) Iterative steering kernel regression [46] (PSNR = 29.80 dB, SSIM = 0.8524). (d) SAIST (PSNR = **31.32** dB, SSIM = **0.8861**).

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

Fig. 10. Image reconstruction performance comparison from a regularly decimated version for *Lena* image. (a) Bicubic interpolation (PSNR = 34.00 dB, SSIM = 0.9146). (b) NEDI [45] (PSNR = 33.94 dB, SSIM = 0.9141). (c) SAI [50] (PSNR = 34.70 dB, SSIM = 0.9174). (d) SAIST (PSNR = 34.27 dB, SSIM = 0.9137).

(a) (b) (c)

Fig. 11. Successful recovery of weak texture patterns by SAIST with a proper annealing schedule. (a) Hat region in the original *Lena* image. (b) Interpolated hat region by SAI suffering from aliasing artifact (similar weakness can be observed in images interpolated by bicubic and NEDI). (c) Interpolated hat region by SAIST ($T_0 = 5.5, \epsilon = 0.05$).

by SVD offers one promising approach in that local variation and nonlocal invariance are respectively characterized in the row and column spaces. It is enlightening to interpret left-multiplying and right-multiplying matrices of SVD - whose energy compaction property has been well known - as the joint sparsification tool and the singular values as the nonlocal variance estimation under a properly chosen dictionary of local bases. From an image modeling perspective, this work is based on two classes of important ideas in the recent literature - dictionary learning (e.g., [57]) and patch clustering (e.g., [3]). What seems a nice surprise is that low-rank approximation achieves nearly-optimal performance for the class of regular edges and textures. Nevertheless, local variation and nonlocal invariance are two sides of the same coin.

B. Optimization: Global Minimum Versus Saddle Point

In our experimental studies, we opted to terminate the SAIST algorithm before reaching the convergence no matter of iterative regularization or deterministic annealing. We have also found that the choices of relaxation parameter in iterative regularization and annealing schedule in deterministic

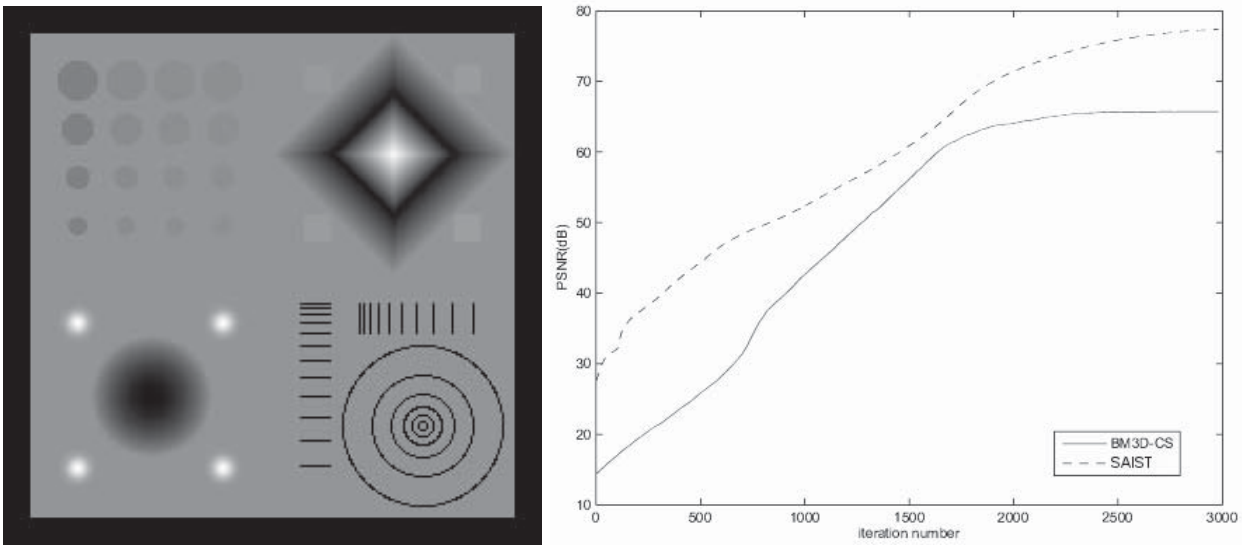


Fig. 12. PSNR performance comparison of MRI reconstruction from 22 radial lines between BM3D-CS [42] and ISVT-DA on *csphantom* (note that for the same image and sampling rate, $l1magic$ can only achieve PSNR = 26.78 dB).

annealing play critical role in the performance of SAIST image denoising and complete algorithms. Moreover, increasingly more empirical evidences have suggested a tantalizing dilemma for image processing researchers - one can often obtain better algorithms by pursuing a nonconvex sparsity optimization [40], [42], [41], [44]; even though the analytical proof or computationally efficient solution to nonconvex problems is often elusive. It is natural to ask: *does nature speak the language of nonconvexity*?

Maybe. Emergent properties in nature (e.g., self-organizing systems [58], dissipative systems [59]) have been widely studied under the context of biological cybernetics and nonequilibrium thermodynamics. If we define image processing as a tool of probing into the mechanism of sensory processing by human vision systems (HVS), it is likely that HVS has developed a strategy of adapting to the emergent properties of natural scenes through evolution and development [60]. In fact, multi-stable perception [61] represents a concrete evidence for the presence of nonconvexity in HVS. As coercively argued by Ashby in [58], the study of a large system built of parts that have "many states of equilibrium" could lead us to identify the "physical basis of the brain's memory". Since nature does not have foresight, we argue that the pursuit of saddle point is physically more plausible than the pursuit of global minimum.

VI. Conclusion

A low-rank (SVD-based) approach toward modeling nonlocal similarity in images was presented, which leads to a conceptually simple image restoration algorithm called Spatially-Adaptive Iterative Singular-value Thresholding (SAIST). The left-multiplying and right-multiplying matrices of SVD jointly characterize the local and nonlocal variations in the row and column spaces of a data matrix. Although excellent energy compaction property of SVD has been well-known, applying it to sparse representation of similar patches gives singular-values a physically plausible interpretation from the bilateral variance estimation point of view. The derived SAIST algorithm can be viewed as a natural extension of soft-thresholding from local wavelet-based to nonlocal SVD-based models. Spatial adaptation of SVD-based models can be achieved by the strategies of iterative regularization and deterministic annealing for noisy and incomplete data respectively. Excellent experimental results of SAIST have been achieved for both image denoising and completion experiments.

To the best of our knowledge, this is the first time that two promising lines of research - namely low-rank method and nonlocal similarity - are unified in a principled way. It might open doors to several promising lines of research. At the modeling level, how do we go beyond translational invariance? How do we accommodate more generic geometric invariance (relevant to object recognition) into the proposed image model? Maybe the boundary between low-level and high-level vision has always been artificial. On the algorithmic side, we believe there is still plenty of room for further improvement - e.g., how to choose patch size and neighborhood size (parameter k in kNN) in a less ad-hoc fashion? how to develop computationally more efficient solution to large-scale (i.e., sample size $\rightarrow \infty$) and parallel (i.e., simultaneously for multiple exemplars) kNN search? We also believe there are plenty of real-world applications (e.g., computational photography, low-light remote sensing, faster MRI) that could benefit from the proposed image model and SAIST algorithm.

Acknowledgment

The authors would like to thank X. Zhang, Xidian University, Xi'an, China, and Prof. L. Zhang, Hong Kong Polytechnic University, Hong Kong, for many helpful discussions.

REFERENCES

[1] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2. 2005, pp. 60–65.

[2] C. Kervrann and J. Boulanger, "Unsupervised patch-based image regularization and representation," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 555–567.

[3] 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.

[4] C. Kervrann and J. Boulanger, "Local adaptivity to variable smoothness for exemplar-based image regularization and representation," *Int. J. Comput. Vis.*, vol. 79, no. 1, pp. 45–69, 2008.

[5] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model. Simul.*, vol. 7, no. 3, pp. 1005–1028, 2008.

[6] A. Elmoataz, O. Lezoray, and S. Bougleux, "Nonlocal discrete regularization on weighted graphs: A framework for image and manifold processing," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1047–1060, Jul. 2008.

[7] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Jun. 2009, pp. 2272–2279.

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

[9] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Mar. 2010, pp. 1791–1798.

[10] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 457–464.

[11] J. Sun and M. F. Tappen, "Learning non-local range Markov random field for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2011, pp. 2745–2752.

[12] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 1–8.

[13] J. Cai, E. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.

[14] E. Candes and Y. Plan, "Matrix completion with noise," *Proc. IEEE*, vol. 98, no. 6, pp. 925–936, Jun. 2010.

[15] A. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[16] G. Stewart, "On the early history of the singular value decomposition," *SIAM Rev.*, vol. 35, no. 4, pp. 551–566, 1993.

[17] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2010, pp. 763–770.

[18] E. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis," *J. ACM*, vol. 58, no. 3, p. 11, 2011.

[19] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.

[20] F. De la Torre and M. Black, "Robust principal component analysis for computer vision," in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, vol. 1. Feb. 2001, pp. 362–369.

[21] M. J. Fadili and J. L. Starck, "EM algorithm for sparse representation-based image inpainting," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2. Sep. 2005, pp. 61–64.

[22] L. Carin, X. Ding, and L. He, "Bayesian robust principal component analysis," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3419–3430, Dec. 2011.

[23] L. Carin, "Nonparametric Bayesian dictionary learning for analysis of noisy and incomplete images," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 130–144, Jan. 2012.

[24] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proc. IEEE*, vol. 86, no. 11, pp. 2210–2239, Nov. 1998.

[25] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.

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

[27] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Royal Stat. Soc., Ser. B, Stat. Methodol.*, vol. 68, no. 1, pp. 49–67, 2006.

[28] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[29] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[30] I. Daubechies, "Orthonormal bases of compactly supported bases," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, Dec. 1988.

[31] X. Li and Y. Zheng, "Patch-based video processing: A variational Bayesian approach," *IEEE Trans. Circuits Systems Video Technol.*, vol. 19, no. 1, pp. 27–40, Jan. 2009.

[32] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. Int. Conf. Image Process.*, 1996, pp. 379–382.

[33] S. Resnick, *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, vol. 10. New York: Springer-Verlag, 2007.

[34] J. Xu and S. Osher, "Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 534–544, Feb. 2007.

[35] M. Charest, M. Elad, and P. Milanfar, "A general iterative regularization framework for image denoising," in *Proc. 40th Annu. Conf. Inf. Sci. Syst.*, 2006, pp. 452–457.

[36] E. Candes and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[37] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. SIGGRAPH*, New Orleans, LA, 2000, pp. 417–424.

[38] O. G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part I: Theory," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 539–554, Mar. 2006.

[39] K. Rose, E. Gurewwitz, and G. Fox, "A deterministic annealing approach to clustering," *Pattern Recognit. Lett.*, vol. 11, no. 9, pp. 589–594, 1990.

[40] O. G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part II: Adaptive algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 555–571, Mar. 2006.

[41] L. Mancera and J. Portilla, "Non-convex sparse optimization through deterministic annealing and applications," in *Proc. Int. Conf. Image Process.*, 2008, pp. 917–920.

[42] K. Egiazarian, A. Foi, and V. Katkovnik, "Compressed sensing image reconstruction via recursive spatially adaptive filtering," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. San Antonio, TX, Sep. 2007, pp. 1–4.

[43] X. Li, "Patch-based nonlocal image interpolation: Algorithms and applications," in *Proc. Local Nonlocal Approx. Image Process.*, 2008, pp. 1–6.

[44] X. Li, "Image recovery via hybrid sparse representations: A deterministic annealing approach," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 953–962, Sep. 2011.

[45] X. Li and M. Orchard, "New edge directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[46] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[47] E. J. Candès, J. K. 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.

[48] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, Sep. 2004.

[49] M. Elad, J. Starck, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmonic Anal.*, vol. 19, no. 3, pp. 340–358, 2005.

[50] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 887–896, Jun. 2008.

[51] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. New York: Academic, 1999.

[52] K. Wilson, "The renormalization group: Critical phenomena and the Kondo problem," *Rev. Modern Phys.*, vol. 47, no. 4, pp. 773–840, 1975.
[53] A. Turing, "The chemical basis of morphogenesis," *Philosoph. Trans. Royal Soc. London, Ser. B, Biol. Sci.*, vol. 237, no. 641, pp. 37–72, 1952.
[54] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*. New York: Springer-Verlag, 1994.
[55] B. B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco, CA: W.H. Freeman, 1982.
[56] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Phys. Rev. Lett.*, vol. 73, pp. 814–817, Aug. 1994.
[57] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
[58] W. Ashby, "Principles of the self-organizing system," *Principles Self-Organizat.*, vol. 6, nos. 1–2, pp. 102–126, 2004.
[59] G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. New York: Wiley, 1977.
[60] O. Schwartz and E. P. Simoncelli, "Natural signal statistics and sensory gain control," *Nature Neurosci.*, vol. 4, no. 8, pp. 819–825, Aug. 2001.
[61] D. Leopold and N. Logothetis, "Multistable phenomena: Changing views in perception," *Trends Cognit. Sci.*, vol. 3, no. 7, pp. 254–264, 1999.

Weisheng Dong received the B.S. degree in electronic engineering from the Huazhong University of Science and Technology, Wuhan, China, and the Ph.D degree in circuits and systems from Xidian University, Xi'an, China, in 2004 and 2010, respectively.

He was a Visiting Student with Microsoft Research Asia, Bejing, China, in 2006. From 2009 to 2010, he was a Research Assistant with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong. In 2010, he joined the School of Electronic Engineering, Xidian University, as a Lecturer, and has been an Associate Professor since 2012. His current research interests include inverse problems in image processing, sparse signal representation, and image compression.

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

Guangming Shi (SM'10) received the B.S. degree in automatic control, the M.S. degree in computer control, and the Ph.D. degree in electronic information technology from Xidian University, Xi'an, China, 1985, 1988, and 2002, respectively.

He joined the School of Electronic Engineering, Xidian University, in 1988. From 1994 to 1996, he was a Research Assistant with the Department of Electronic Engineering, University of Hong Kong. Since 2003, he has been a Professor with the School of Electronic Engineering, Xidian University, and, since 2004, the Head of the National Instruction Base of Electrician & Electronic (NIBEE). In 2004, he had studied with the Department of Electronic Engineering, University of Illinois at Urbana-Champaign, Urbana. Currently, he is the Deputy Director of the School of Electronic Engineering, Xidian University, and the Academic Leader of the subject of circuits and systems. He has authored or co-authored over 60 research papers. His current research interests include compressed sensing, theory and design of multirate filter banks, image denoising, low-bit-rate image/video coding and implementation of algorithms for intelligent signal processing (using DSP&FPGA).

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

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

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