Image Deblurring with Low-rank Approximation Structured Sparse Representation

Weisheng Dong*, Guangming Shi*, and Xin Li†
*Key Laboratory of Intelligent Perception and Image Understanding (Chinese Ministry of Education) School of Electronic Engineering, Xidian University, China
E-mail: wsdong@mail.xidian.edu.cn
†Lane Dept. of Comp. Sci. and Elec. Engr., West Virginia University, Morgantown, WV 26506

Abstract— **In recent years sparse representation model (SRM) based image deblurring approaches have shown promising image deblurring results. However, since most of the current SRMs don't utilize the spatial correlations between the nonzero sparse coefficients, the SRM-based image deblurring methods often fail to faithfully recover sharp image edges. In this paper, a structured SRM is employed to exploit the local and nonlocal spatial correlation between the sparse codes. The connection between the structured SRM and the low-rank approximation model has also been exploited. An effective image deblurring algorithm using the patch-based structured SRM is then proposed. Experimental results demonstrate the improvements of the proposed deblurring method over current state-of-the-art image deblurring methods.**

I. INTRODUCTION

Image deblurring aiming to recover a clean and sharp image from a noisy and blurred observation can be modeled as

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

where $\mathbf{H} \in \Re^{N\times N}$ and $\boldsymbol{v} \in \Re^{N}$ denote a blurring matrix and Gaussian noise, respectively. Recovering the original image $\boldsymbol{x} \in \Re^{N}$ from the observation $\boldsymbol{y} \in \Re^{N}$ is a typical ill-posed inverse problem. Due to its ill-posed nature of the deblurring problem, regularization techniques that incorporate prior knowledge of natural images are required for finding a reasonable solution. The regularized solution of the deblurring problem can be obtained by solving the following optimization problem:

$$\boldsymbol{x} = \arg\min_{\boldsymbol{x}} \| \boldsymbol{y} - \mathbf{H}\boldsymbol{x} \|_2^2 + \lambda J(\boldsymbol{x}) , \tag{2}$$

where $J(\boldsymbol{x})$ denotes a regularization term incorporating the prior knowledge of the unknown image $\boldsymbol{x}$, and λ is a constant balancing the tradeoff between the data fidelity term and the regularization term. In the regularization based deblurring approaches, the construction of an effective regularization term is of great importance. Classic regularizers, including quadratic Tikhonov regularization and the well-known total variation (TV) regularization, are effective in removing the noise artifacts but tend to produce over smoothed results.

In recent years the emerging of the compressive sensing theory [1] has boosted a flurry of research on the sparse representation model (SRM) based image restoration [2-8]. The SRM assumes that natural images can be sparsely represented under a basis or an overcomplete dictionary $\mathbf{U}$, i.e., $\boldsymbol{x} \approx \mathbf{U}\boldsymbol{\alpha}$, $\|\boldsymbol{\alpha}\|_1 \le T$. The sparsity-regularized deblurring approach can be formulated as

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

which can be efficiently solved by the iterative shrinkage algorithm [8]. The deblurred image is then obtained by $\hat{\boldsymbol{x}} = \mathbf{U}\hat{\boldsymbol{\alpha}}$. Since most of natural images allow sparsely representations under common bases (e.g., DCT, wavelet, TV), the SRM-based deblurring approaches have shown promising results [3, 6-7]. By adapting the dictionaries to the image content, the deblurring results can be further improved [6-7].

In Eq. (3), the l_1-sparsity penalty term assumes that the coefficients are identical independent distributed (i.i.d), disregarding the correlation between the sparse coefficients. The plain l_1-sparsity model has been shown to be unstable, especially when the observed image is degraded (e.g., noisy and blurry). In such image restoration scenario, the sparse model may fail to recover accurate sparse codes, leading to unsatisfied reconstructed images. Recently, efforts have been made to leverage the structural correlation in the sparse model. In [7], centralized sparse representation (CSR) model has been proposed to utilize the nonlocal redundancies, leading to state-of-the-art image deblurring results. Instead of sparsely coding each patch individually, simultaneous sparse coding (SSC) techniques [4] code a set of patches simultaneously to encourage the alignment of the sparse codes. By exploiting the nonlocal self-similarity, the SSC techniques have led to state-of-the-art denoising results [4]. However, despite its effectiveness, the optimization of the learned SSC model is in general difficult, which requires alternative optimization of the sparse codes and the dictionary. Moreover, deeper reasoning of the SSC model is needed to better understand its effectiveness.

On the other hand, as an effective tool for high dimensional data modeling, the low-rank approximation (LA) has attracted more and more attentions. The applying of singular value decomposition (SVD) based low-rank approximation suggests an intrinsic connection between the SSC model and the low-rank approximation. In our previous work [5], such connection between SSC and LA has been exploited and a principle optimization approach for SSC has been proposed using the SVD thresholding, leading to state-of-the-art denoising results. In this paper, an effective SSC-based deblurring method is proposed using the *patch* based low-rank approximation structured sparse coding (LASSC) algorithm. Experimental results show that the proposed LASSC deblurring method can achieve convincing improvements over current state-of-the-art deblurring methods.

II. STRUCTURED SPARSE CODING VIA LOW-RANK APPROXIMATION

Instead of sparsely coding each patch individually, the SSC approach [4] simultaneously codes a set of patches. Mathematically, the SSC model can be formulated as

$$(\mathbf{A},\mathbf{U})=\arg\min_{A}\|\mathbf{X}-\mathbf{UA}\|_2^2+\tau J(\mathbf{U}), \quad (4)$$

where $\mathbf{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,...,\boldsymbol{x}_m]\in\Re^{n\times m}$, $\mathbf{A}=[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,...,\boldsymbol{\alpha}_m]\in\Re^{n\times m}$, denote dataset and sparse coefficients matrix, respectively, $\mathbf{U}\in\Re^{n\times K}$ ($n\leq K$) denotes the dictionary, τ is a regularization parameter, and $J(\mathbf{A})$ denotes a matrix norm defined as

$$J(\mathbf{A})=\sum_{i=1}^{n}\omega_i\|\boldsymbol{\alpha}^i\|_q^p, \quad (5)$$

where $\boldsymbol{\alpha}^i$ denotes the coefficient vector of the i^{th} row of matrix $\mathbf{A}$ and ω_i is the weight. The data set $\mathbf{X}$ is constructed by grouping the nonlocal similar patches. For a convex optimization, the $l_{1,2}$ norm is often used. Since the patches in $\mathbf{X}$ share similar edge structures, overcomplete dictionary learning approaches are not needed, a compact dictionary (e.g., the principle component analysis (PCA) dictionary) learned from $\mathbf{X}$ is good enough to sparsely represent $\mathbf{X}$. The local PCA dictionaries have been successfully used in [6-7]. In this paper, we use the local PCA dictionary learned from $\mathbf{X}$. Another important issue is the selection of the weights ω_i. The weights ω_i play a similar role to that of the weighted l_1-norm in reweighted sparse coding [9]. In [4], all the weights were set to 1, which were obviously non-optimal.

In this section we briefly review the low-rank approximation structured sparse coding (LASSC) algorithm presented in our previous work [5], and the deblurring algorithm using LASSC will be presented in Sec. III.

A. Low-rank approximation SSC

For convexity $l_{1,2}$ norm is used in the structured matrix norm. A key observation is that the $l_{1,2}$ norm in Eq. (5) can be rewritten as

$$J(\mathbf{A})=\sum_{i=1}^{n}\omega_i\|\boldsymbol{\alpha}^i\|_2=\sum_{i=1}^{n}\omega_i\sqrt{\alpha_{i,1}^2+\alpha_{i,2}^2+\cdots+\alpha_{i,m}^2}=\sqrt{m}\sum_{i=1}^{n}\omega_i\sigma_i \quad (6)$$

where σ_i denotes the standard derivation of the coefficients in the i^{th} row of $\mathbf{A}$. Therefore, the structured sparsity can be interpreted as the variance sparsity. Substituting Eq. (6) into Eq. (4), we obtain the following objective function

$$(\mathbf{A},\mathbf{U})=\arg\min_{A}\|\mathbf{X}-\mathbf{UA}\|_2^2+\tau\sqrt{m}\sum_{i=1}^{n}\omega_i\sigma_i. \quad (7)$$

To facilitate the SSC from the variance sparsity perspective, we further decomposite the matrix $\mathbf{A}$ into two components, as follows

$$\mathbf{A}=\mathbf{\Sigma}\mathbf{Q}^T, \quad (8)$$

where $\mathbf{\Sigma}=diag\{\sqrt{m}\sigma_1,\sqrt{m}\sigma_2,\cdots,\sqrt{m}\sigma_r\}$ is a diagonal matrix, where r=max$\{n,m\}$, and each column of $\mathbf{Q}\in\Re^{m\times r}$ is defined as $\boldsymbol{q}_i=(\boldsymbol{\alpha}^i)^T/(\sqrt{m}\sigma_i)$. Substituting Eq. (8) into Eq. (7), we obtain

$$(\mathbf{U},\mathbf{\Sigma},\mathbf{Q})=\arg\min_{A}\|\mathbf{X}-\mathbf{U\Sigma Q}^T\|_2^2+\lambda\sqrt{m}\sum_{i=1}^{n}\omega_i\sigma_i. \quad (9)$$

Compared to the original SSC problem in Eq. (4), Eq. (9) involves another unknown dictionary $\mathbf{Q}$ that plays the role of sparsifying the coefficient matrix. However, the connection between the data term and the regularization term by $\boldsymbol{\sigma}_i$ makes the optimization of Eq. (9) decomposable. The structured sparsity regularizer of Eq. (9) is actually a weighted l_1-sparsity regularizer imposed on the field of variance (standard deviation). Thus, Eq. (9) can be solved by an iterative thresholding algorithm. Moreover, the right and left multiplying matrixes on two sides of $\mathbf{\Sigma}$ show striking resemblance to a well-known factorization tool known as SVD. Such resemblance motivates us to exploit the connection between the SSC and the recent developed low-rank approximation (LA) methods [10-12].

Let $\mathbf{X}=\mathbf{U\Lambda V}^T$ be the SVD of the dataset $\mathbf{X}$, where $\mathbf{\Lambda}=diag\{\lambda_1,\lambda_2,\cdots,\lambda_r\}$ and $\mathbf{V}\in\Re^{m\times r}$. Note that $\boldsymbol{\alpha}^i=\lambda_i\boldsymbol{v}_i^T$, where $\boldsymbol{v}_i$ is the i^{th} column of matrix $\mathbf{V}$. Then we have

$$m\sigma_i^2=\|\boldsymbol{\alpha}^i\|_2^2=\|\lambda_i\boldsymbol{v}_i^T\|_2^2=\lambda_i^2, \quad (10)$$

where the first equation comes from Eq. (6) and the second equation comes from the unitary of the matrix $\mathbf{V}$. Comparing Eq. (10) and Eq. (9), we can see that they are equivalent, and Eq. (9) can be rewritten as

$$(\mathbf{U},\mathbf{\Lambda},\mathbf{V})=\arg\min_{\mathbf{U},\mathbf{\Lambda},\mathbf{V}}\|\mathbf{X}-\mathbf{U\Lambda V}^T\|_2^2+\tau\sum_{i=1}^{n}\omega_i\lambda_i. \quad (11)$$

This is a weighted LA optimization problem. When ω_i are set to 1, Eq. (11) is a typical LA problem. The equivalence between LA and the SSC allows us to solve the SSC optimization problem of Eq. (7) or Eq. (11) by efficient low-rank solvers, such as the SVD thresholding algorithm [12] that admits the following closed form solution:

$$\begin{cases}(\mathbf{U},\mathbf{\Lambda},\mathbf{V})=\text{svd}(\mathbf{X})\\ \hat{\lambda}_i=S_{\tau\cdot\omega_i}(\lambda_i)\end{cases}, \quad (12)$$

where $S_{\tau\cdot\omega_i}(\cdot)$ denotes the soft thresholding operator with threshold $\tau\cdot\omega_i$. Then we can see that the dictionary $\mathbf{U}$ and the sparse code $\mathbf{A}$=$\mathbf{\Lambda V}^T$ can be jointly optimized by the principled algorithm.

B. Adaptive iterative SVD thresholding

An important issue of the SSC is the selection of the weights and the regularization parameters τ. When ω_i are set to 1, then Eq. (11) is a standard LA problem. By proper selection of the weights, we may further enhance the sparsity and consequently improve the signal reconstruction. Similar reweighted l_1-sparse coding has been successfully used for image reconstruction [6, 9]. Here we propose to select the weights toward a MAP estimation of the variance.

According to Eq. (10), Eq. (11) can be rewritten as

$$(\mathbf{U},\mathbf{\Lambda},\mathbf{V}) = \arg\min_{\mathbf{U},\mathbf{\Lambda},\mathbf{V}} \|\mathbf{X}-\mathbf{U}\mathbf{\Lambda}\mathbf{V}^T\|_2^2 + \tau\sqrt{m}\sum_{i=1}^{n}\omega_i\sigma_i , \quad (13)$$

from which we can see that the joint sparsity of the sparse codes $\mathbf{A}$ is equivalent to the l_1-norm variance. Previous studies on l_1-norm sparse coding have shown that the l_1-sparse model is equivalent to the MAP estimator using a Laplaican distribution [6]. By estimating the parameters of the local Laplacian distribution, the regularization parameters can be adaptively selected [6]. Such connection can also be extended from the magnitude domain to the variance domain, leading to the following selection of the weights:

$$\tau = 2\sqrt{2}\sigma_w^2,\ \omega_i = 1/\theta_i , \quad (14)$$

where σ_w^2 denotes the variance of the additive Gaussian noise, θ_i denote standard deviation of the variance σ_i. For the details of the deviation, please refer to [5]. In practice, it is difficult to estimate θ_i due to the lack of the sufficient samples of σ_i. Here we adopt an one sample maximum likelihood estimation of θ_i:

$$\theta = \sqrt{\max(\tilde{\sigma}_i^2 - \sigma_w^2, 0)} , \quad (15)$$

where $\tilde{\sigma}_i^2$ is an estimate of σ_i, which can be computed as $\tilde{\sigma}_i^2 = \tilde{\lambda}_i^2 / m$.

III. PROPOSED IMAGE DEBLURRING ALGORITHM

In the deblurring problem the observed image is given by $\boldsymbol{y}$=$\mathbf{H}\boldsymbol{x}$+$\boldsymbol{n}$, where $\mathbf{H}$ is a blurring kernel assuming to be known. Note that the observation model is a general model and $\mathbf{H}$ can represent other degradation operators, such as compressive sensing and super-resolution. Now the question is how to use the *patch*-based SSC model for the *whole* image deblurring. The basic idea is to enforce the structured sparsity over the grouped nonlocal similar patches for each extracted exemplar patch, subjected to the constraint of the observation model. The proposed objective function of the structured sparsity regularized image deblurring method can be formulated as

$$(\hat{\boldsymbol{x}},\hat{\mathbf{U}}_i,\hat{\mathbf{\Lambda}}_i,\hat{\mathbf{V}}_i) = \arg\min_{\boldsymbol{x},\mathbf{U}_i,\mathbf{\Lambda}_i,\mathbf{V}_i} \|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|_2^2 + \beta\sum_i\sum_{j\in G_i}\|\mathbf{R}_j\boldsymbol{x}_j - \mathbf{U}_i\lambda_j\boldsymbol{v}_j^T\|_2^2 + \tau\sum_i\sum_{j\in G_i}\omega_j\lambda_j , \quad (16)$$

where G_i denotes the group of the patches similar to exemplar patch $\boldsymbol{x}_i$, λ_j and $\boldsymbol{v}_j$ denotes the j^{th} singular value and j^{th} column of matrix $\mathbf{V}_j$, respectively. Eq. (16) can be solved by alternatively optimizing $\boldsymbol{x}$ and the sets of $\{\mathbf{U}_i\}$, $\{\mathbf{\Lambda}_i\}$ and $\{\mathbf{V}_i\}$. That is, for fixed $\mathbf{X}_i^{(k)} = \mathbf{U}_i^{(k)}\mathbf{\Lambda}_i^{(k)}(\mathbf{V}_i^{(k)})^T$ obtained at k^{th} iteration, the whole image can be updated using the set of reconstructed patches $\mathbf{X}_i^{(k)}$ by solving the following sub-optimization problem:

$$\boldsymbol{x}^{(k+1)} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y}-\mathbf{H}\boldsymbol{x}\|_2^2 + \beta\sum_i\sum_{j\in G_i}\|\mathbf{R}_j\boldsymbol{x}_j - \mathbf{U}_i^{(k)}\lambda_j^{(k)}(\boldsymbol{v}_j^{(k)})^T\|_2^2 , \quad (17)$$

which is a quadratic optimization problem that admits a closed form solution, i.e.,

$$\boldsymbol{x}^{(k+1)} = \frac{[\mathbf{H}^T\boldsymbol{y} + \beta\sum_i\sum_{j\in G_i}\mathbf{R}_j^T\mathbf{U}_i^{(k)}\lambda_j^{(k)}(\boldsymbol{v}_j^{(k)})^T]}{(\mathbf{H}^T\mathbf{H} + \beta\sum_i\sum_{j\in G_i}\mathbf{R}_j^T\mathbf{R}_j)} , \quad (18)$$

Since the matrix to be inverted is very large, in practice the conjugate gradient (CG) algorithm is employed to compute Eq. (18). For a fixed estimate of the deblurred image $\boldsymbol{x}^{(k)}$, the dataset $\mathbf{X}_i$ of each exemplar patch can be updated by solving a reweighted low-rank approximation problem:

$$(\mathbf{U}_i^{(k)},\mathbf{\Lambda}_i^{(k)},\mathbf{V}_i^{(k)}) = \arg\min_{\mathbf{U}_i,\mathbf{\Lambda}_i,\mathbf{V}_i} \sum_i\sum_{j\in G_i}\|\mathbf{R}_j\boldsymbol{x}^{(k)} - \mathbf{U}_i\lambda_j\boldsymbol{v}_j^T\|_2^2 + \frac{\tau}{\beta}\sum_i\sum_{j\in G_i}\omega_j\lambda_j , \quad (19)$$

which also admits a closed-form solution, i.e.,

$$\begin{cases}(\mathbf{U}_i^{(k+1)},\mathbf{\Lambda}_i^{(k+1/2)},\mathbf{V}_i^{(k+1)}) = \mathrm{svd}(\mathbf{X}_i^{(k)}) \\ \lambda_j^{(k+1)} = S_{\tau\cdot\omega_j/\beta}(\lambda_j^{(k+1/2)})\end{cases} , \quad (20)$$

where $\mathbf{X}_i^{(k)}$ decomposed of the similar patches of exemplar $\boldsymbol{x}_i^{(k)}$. The dataset $\mathbf{X}_i$ is then updated as $\mathbf{X}_i^{(k+1)} = \mathbf{U}_i^{(k+1)}\mathbf{\Lambda}_i^{(k+1)}(\mathbf{V}_i^{(k+1)})^T$. The above alternative optimization process can be iterated until convergence. The overall algorithm is summarized in **Algorithm 1**.

Algorithm 1

---

1. Initialization:
   - (a) Set the initial estimate as $\hat{\boldsymbol{x}} = \boldsymbol{y}$;
   - (b) Patch clustering: find the k-NN for each exemplar patch and create data matrix $\mathbf{X}_i$ for each cluster;
2. Iterate on k =1,2,…, K
   - (a) SVD for each data matrix $\mathbf{X}_i$: $(\mathbf{U}_i^{(k+1)},\mathbf{\Lambda}_i^{(k+1/2)},\mathbf{V}_i^{(k+1)}) = \mathrm{svd}(\mathbf{X}_i^{(k)})$;
   - (b) Singular value thresholding using Eq. (20);
   - (c) Image update: update the deblurred image $\boldsymbol{x}^{(k+1)}$ by computing Eq. (18);
   - (d) Threshold update: compute ω_i using Eqs. (14) and (15);
   - (e) If *mod*(*k, K*$_0$)=0 update the patch clustering;

In **Algorithm 1**, we update the patch clustering in every K_0 (K_0=20) iterations to save computation complexity. In general, **Algorithm 1** converges after 240 iterations.

IV. EXPERIMENTAL RESULTS

To verify the deblurring performance of the proposed method, we implemented the proposed low-rank approximation structured sparse coding (LASSC) based deblurring method and conduct experiments on several natural images. The basic parameters of **Algorithm 1** are as follows: the patch size is 6×6; total 41 similar patches are selected for each chosen exemplar; β=0.5. Some recently developed non-blind deblurring methods are selected for comparison study, including the constrained TV method (denoted by FISTA) [13], the well-known BM3D method [3], and our previous CSR deblurring method [7]. Note that the CSR method is one of the current state-of-the-art non-blind deblurring methods.

The blurred images are generated by applying a blur kernel to the original test images, and additive Gaussian noise with standard deviation $\sigma_w = \sqrt{2}$ is also added to the blurred images. Two blur kernels, i.e., 9×9 uniform blur kernel and 2D Gaussian blur kernel (nontruncated) with standard deviation 1.6 are used for simulation. For color images, we only applied the deblurring method to the luminance component. The PSNR results are reported in Tables 1 and 2, from which we can see that the proposed LASSC method significantly outperforms the FISTA and BM3D methods, and also performs better than CSR method on most test images. The PSNR gain over CSR can be up to 0.68 dB. The visual comparison of the deblurring methods is shown in Figs. 1~2. It can be seen that the proposed LASSC method is very effective in recovering sharp and clean image structures. By contrast, other deblurring methods tend to generate artifacts around edges.

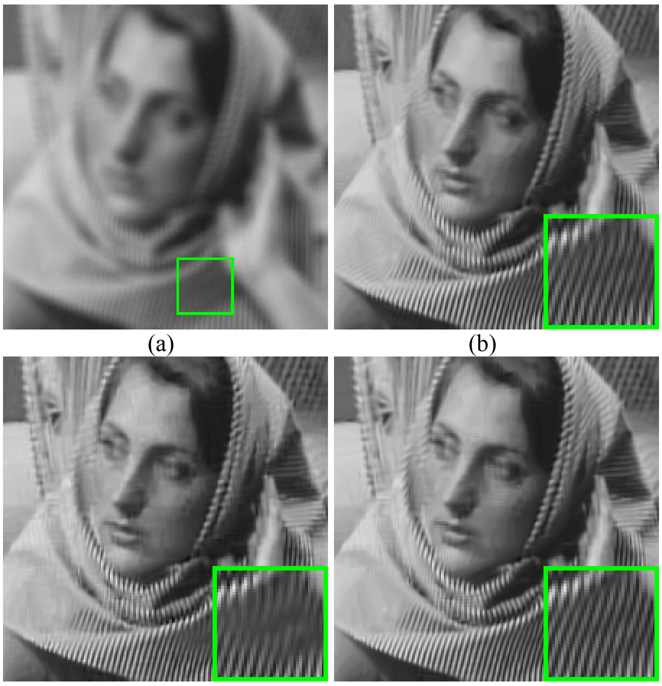
(a) (b) (c) (d)

Fig. 1 Deblurring performnce comparison for Barbara image (9×9 uniform blur, $\sigma_n = \sqrt{2}$). (a) Noisy and blurred; (b) BM3D [3] (PSNR= 27.99 dB); (c) CSR [7] (PSNR=27.93 dB); (d) Proposed LASSC (PSNR=**28.40** dB).

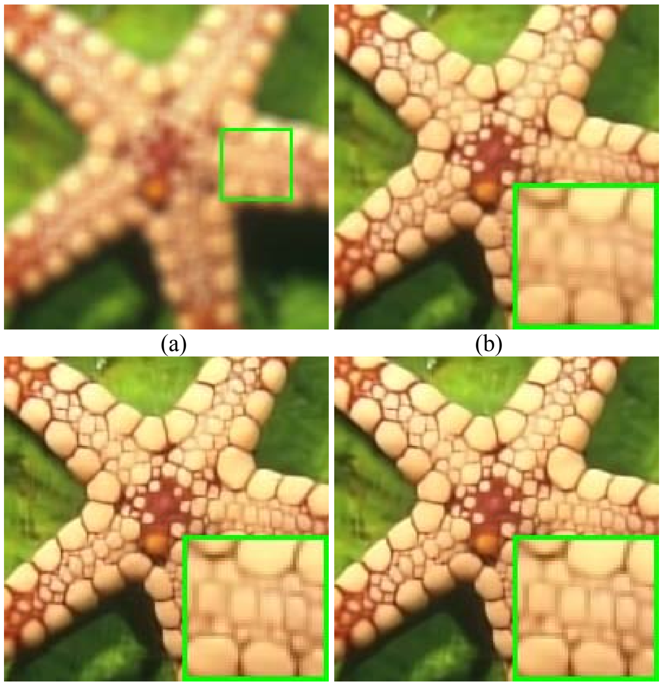
(a) (b) (c) (d)

Fig. 2 Deblurring performnce comparison for *Starfish* image (9×9 uniform blur, $\sigma_n = \sqrt{2}$). (a) Noisy and blurred; (b) BM3D [3] (PSNR=28.61 dB); (c) CSR [7] (PSNR=30.30 dB); (d) Proposed LASSC (PSNR=**30.72** dB).

Table 1 PSNR results of the test methods (9×9 uniform blur, $\sigma_n = \sqrt{2}$)

Images	*Butterfly*	*Parrot*	*Starfish*	*Barbara*	*Leaves*
TV [13]	28.37	29.11	27.75	25.75	26.49
BM3D [3]	27.21	30.50	28.61	27.99	27.45
CSR [7]	29.75	32.09	30.30	27.93	29.97
LASSC	**30.16**	**32.33**	**30.72**	**28.40**	**30.65**

Table 2 PSNR results of the test methods (Gaussian blur, $\sigma_n = \sqrt{2}$)

Images	*Butterfly*	*Parrot*	*Starfish*	*Barbara*	*Leaves*
TV [13]	30.36	31.27	29.65	25.03	29.36
BM3D [3]	29.01	32.22	30.71	28.19	29.67
CSR [7]	30.75	**33.44**	**32.31**	27.81	31.44
LASSC	**30.79**	33.41	32.29	**28.44**	**31.74**

V. CONCLUSIONS

In this paper we present a novel low-rank approximation based structured sparse representation (LASSC) model for image deblurring. The intrinsic connection between the structured sparse coding and the low-rank approximation has been expoilted to develop an efficient singular value thresholding algorithm for structured sparse coding. An effective image deblurring method using the patch-based LASSC model is then presented. Experimental results show that the proposed LASSC image deblurring produces the state-of-the-art image deblurring results.

ACKNOWLEDGEMENT

This work was supported by the Natural Science Foundation of China under (Grants 61100154, 61033004, 61070138, and Grant 61072104), in part by Major State Basic Research Development Program of China (973 Program) (No. 2013CB329402), in part by the Fundamental Research Funds of the Central Universities of China under Grant K50510020003, in part by NSF-CCF-0914353 and NSF-ECCS-0968730.

REFERENCES

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

[2] 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.

[3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series,* vol. 6812, 2008.

[4] J. Mairal, F. Bach, J. Ponce, G. Sapiro and A. Zisserman, "Non-Local Sparse Models for Image Restoration," in *Proc. IEEE International Conference on Computer Vision*, Tokyo, Japan, 2009.

[5] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation : a low-rank approach, " *IEEE Trans. On Image Processing*, to appear, 2012.

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

[7] W. Dong, L. Zhang and G. Shi, "Centralized sparse representation for image restoration," in *Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, 2011.

[8] I. Daubechies, M. Defriese, and C. DeMol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math*., vol.57, pp.1413-1457, 2004.

[9] E. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted l_1 minimization," Journal of Fourier Analysis and Applications, vol. 14, pp. 877-905, 2008.

[10] E. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis", *Journal of the ACM*, vol. 58, no. 3, 2011.

[11] E. Candes and Y. Plan, "Matrix completion with noise," *Proceeding of the IEEE*, vol. 98, no. 6, pp. 925-936, 2010.

[12] J. Cai, E. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, 2010.

[13] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. On Image Process.,* vol. 18, no. 11, pp. 2419-2434, Nov. 2009.