

An improved nonlocal sparse regularization-based image deblurring via novel similarity criteria

Nannan Wang¹, Wenxuan Shi^{2,3} , Ci'en Fan¹ and Lian Zou¹

Abstract

Image deblurring is a challenging problem in image processing, which aims to reconstruct an original high-quality image from its blurred measurement caused by various factors, for example, imperfect focusing caused by the imaging system or different depths of scene appearing commonly in our daily photos. Recently, sparse representation whose basic idea is to code an image patch as a linear combination of a few atoms chosen out from an over-complete dictionary has shown uplifting results in image deblurring. Based on this and another heart-stirring property called nonlocal self-similarity, some researchers have developed nonlocal sparse regularization models to unify the local sparsity and the nonlocal self-similarity into a variational framework for image deblurring. In such models, the similarity evaluation for searching similar image patches is indispensable and influential in deblurring performance. Though the traditional Euclidean distance is generally a choice as a similarity metric, its application might give rise to inferior performance since it fails to capture the intrinsic structure of image patches. Consequently, in this article, based on structural similarity index and principal component analysis, we propose the nonlocal sparse regularization-based image deblurring with novel similarity criteria called structural similarity distance and principal component analysis-subspace Euclidean distance to improve the accuracy of deblurring. The structural similarity index is commonly used for assessing perceptual image quality, and principal component analysis is pervasively used in pattern recognition and dimensionality reduction. In our comprehensive experiments, the nonlocal sparse regularization-based image deblurring with our novel similarity criteria has achieved higher peak signal-to-noise and favorable consistency with subjective vision perception compared with state-of-the-art deblurring algorithms.

Keywords

Image deblurring, sparse representation, nonlocal self-similarity, principal component analysis, structural similarity index

Date received: 8 January 2018; accepted: 23 May 2018

Topic: Vision Systems

Topic Editor: Antonio Fernandez-Caballero

Associate Editor: Antonios Gasteratos

Introduction

Image deblurring has many applications from astronomical imaging and remote sensing imaging to medical imaging. It has drawn attention around the world for decades and has been regarded as a worthy research topic in image processing so far. If we can acquire the degraded image \mathbf{Y}_{im} of size $\sqrt{N} \times \sqrt{N}$ pixels, then the task of image deblurring is to

¹Electronic Information School, Wuhan University, Wuhan, China

²Remote Sensing and Information Engineering School, Wuhan University, Wuhan, China

³Collaborative Innovation Center for Geospatial Technology, Wuhan University, Wuhan, China

Corresponding author:

Wenxuan Shi, Remote Sensing and Information Engineering School, Wuhan University, Wuhan 430072, Hubei, China.

Email: shiwx@whu.edu.cn



recover the unknown clean image \mathbf{X}_{im} based on the following degradation model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + n \quad (1)$$

where $\mathbf{x} \in \mathcal{R}^N$ and $\mathbf{y} \in \mathcal{R}^N$ are column stacked vector versions of \mathbf{X}_{im} and \mathbf{Y}_{im} , respectively, that is, $\mathbf{x} = [x_{(1)}, x_{(2)}, \dots, x_{(N)}]^T$ and $\mathbf{y} = [y_{(1)}, y_{(2)}, \dots, y_{(N)}]^T$; n is the additive Gaussian noise, that is, $n \sim N(0, \sigma^2)$; and \mathbf{H} is an irreversible blur matrix involved with a point spread function (PSF) kernel. Whether \mathbf{H} is known or not, image deblurring can be divided into two types, non-blind and blind deblurring. As for blind deblurring, when the blur matrix \mathbf{H} is estimated by using some method,^{1,2} the strategies used in non-blind deblurring are appropriate for blind deblurring. For a compact expression hereinafter, we use column vectors \mathbf{x} and \mathbf{y} to refer to the clean image and the blurred image, respectively.

Obviously, image deblurring is an ill-posed linear inverse problem³; in other words, with the degraded image \mathbf{y} and the known blur matrix \mathbf{H} , there can be more than one clean image \mathbf{x} that satisfies equation (1). Hence, traditional deblurring approaches in terms of deconvolution^{4,5} by inverting the blur matrix \mathbf{H} straightforwardly often introduce severe ringing artifacts controlled by the additive noise n .

To alleviate this situation, the regularization models, which try to incorporate both the observation model and the prior knowledge of the clean image as a regularization term into a variational formulation, have been widely investigated and adopted for image deblurring. The regularization models are generally formulated as the following minimization problem

$$\arg \min_{\mathbf{x}} \{ \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \lambda R(\mathbf{x}) \} \quad (2)$$

where $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$ is a quadratic data-fidelity term, $R(\mathbf{x})$ is called regularization term modeling the prior knowledge, and constant λ named the regularization parameter can balance between $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$ and $R(\mathbf{x})$ with a proper value.

Because the design of effective regularization terms is at the core of image deblurring, some classical regularization terms have been designed, for example, quadratic Tikhonov regularization,⁶ Mumford–Shah regularization,⁷ wavelet regularization,⁸ or total variation (TV) regularization.⁹ Even though various deblurring methods are rising and developed, the regularization models still show their extraordinary charm in image deblurring. This has motivated many outstanding investigators to start a long process of exploration to find excellent image priors as regularization terms.

In the exploration process, sparsity as one of the most significant properties of natural images has gradually come into researchers' sight in image deblurring. As a result, sparse representation-based regularization models^{10–12} have been promoted rapidly and have achieved great

success in image deblurring. These models follow an assumption that each patch of an image can be precisely represented by a few elements from a basis set called a dictionary. Instead of traditional analytically designed dictionaries¹³ based on transform bases, such as discrete cosine transform, wavelet, or curvelet, the dictionaries learned from example image patches^{10,14} adapt better to local image structures and achieve better deblurring performance.

Mathematically, for a given dictionary \mathbf{D} , a signal $\mathbf{x} \in \mathcal{R}^N$ can be sparsely represented as $\mathbf{x} \approx \mathbf{D}\boldsymbol{\alpha}$ by solving an l_0 -minimization problem, showed as $\arg \min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0$, s.t. $\|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2 < \delta$, where δ is a small constant controlling the approximation error. For efficient convex optimization,¹⁵ the NP-hard l_0 -minimization is often relaxed to its counterpart l_1 -minimization and we formulate this l_1 -minimization as the following regularization form

$$\boldsymbol{\alpha}_x = \arg \min_{\boldsymbol{\alpha}} \{ \|\mathbf{x} - \mathbf{D}\boldsymbol{\alpha}\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1 \} \quad (3)$$

However, in equation (3), only the local sparsity is considered; in other words, each exemplar patch is usually regarded to be independent in dictionary learning and sparse coding. Considering this problem, lately, when nonlocal self-similarity which describes the repetitiveness existed in textures and structures among nonlocal image patches is well-known, a series of models regularized by nonlocal self-similarity^{16–18} are emerging.

More recently, the so-called nonlocal sparse regularization models^{19–22} which combine the local sparsity and the nonlocal self-similarity into a unified framework are becoming more and more popular in image deblurring. Among these models, a nonlocally centralized sparse representation (NCSR) model²⁰ proposed by Dong et al. is superior to others. In this article, our nonlocal sparse regularization-based image deblurring via novel similarity criteria is realized under the NCSR framework.

In the NCSR model, first, many overlapped image patches with similar structures are included in the same cluster using a k-means clustering algorithm.^{23,24} Then, Zhang et al. exploit the principal component analysis (PCA) technique²⁵ and view each cluster as the basic unit to learn a series of compact subdictionaries. Next, during the sparse coding phase, the best subdictionary that is most fitted to code a given patch is selected to obtain the sparse coding coefficients. Last, the sparse coding noise (SCN), defined as the deviation between sparse coding coefficients of the blurred image and a good estimation of sparse coding coefficients of the clean image based on a nonlocal means (NLM) method,²⁶ is suppressed to improve the performance of image deblurring.

Analyzing the entire process in the NCSR modeling, we can discover that the similarity evaluation for searching similar image patches is involved not only in k-means clustering for dictionary learning and subdictionary selection for each patch but also in estimation of the sparse coding

coefficients by the principle of NLM. On this account, the accuracy of evaluating the similarity of image patch pairs is a vital factor for satisfactory deblurring quality. The Euclidean distance as a similarity metric is a usual choice; however, the Euclidean distance is too simple to evaluate the similarity among image patches precisely without capturing the intrinsic structure of patches, which has limited the performance of deblurring. To suppress this problem, in this article, we propose to adopt structural similarity (SSIM)²⁷ index and the Euclidean distance in a lower dimensional space made by PCA as new similarity criteria. Extensive experiments are conducted on image deblurring, which demonstrate that our nonlocal sparse regularization-based image deblurring via novel similarity criteria under the NCSR framework can outperform significantly the original NCSR model and other state-of-the-art methods both in visual effect and quantitative evaluation via peak signal-to-noise (PSNR), SSIM, and feature similarity (FSIM).²⁸

With the rapid rise of big data and deep learning in last few years, more and more scholars attempt to utilize deep learning in image deblurring and many worthy works^{29–32} have been done. Though this kind of deblurring methods can deblur images to a certain degree, they are more likely to ignore high frequency information in images and the fine image structures.

Therefore, the following sections are presented to put forward our solutions. “Ideas on our novel similarity criteria” section develops the ideas in our proposed SSIM distance and PCA-subspace Euclidean distance. “Image deblurring via our proposed novel similarity criteria under the NCSR framework” section introduces image deblurring via our proposed novel similarity criteria under the NCSR framework in detail. “Experimental results” section presents extensive experimental results and “Conclusion and future work” section summarizes this article.

Ideas on our novel similarity criteria

For an image \mathbf{X}_{im} of size $\sqrt{N} \times \sqrt{N}$ pixels and an image patch \mathbf{x}_{im_i} of size $\sqrt{n} \times \sqrt{n}$ pixels extracted from \mathbf{X}_{im} at location i , let $\mathbf{x} \in \mathcal{R}^N$, $\mathbf{x}_i \in \mathcal{R}^n$ denote column vectors obtained by stacking the columns of \mathbf{X}_{im} and \mathbf{x}_{im_i} , respectively, that is, $\mathbf{x} = [x_{(1)}, x_{(2)}, \dots, x_{(N)}]^T$, $\mathbf{x}_i = [x_{i(1)}, x_{i(2)}, \dots, x_{i(n)}]^T$, and $\{\mathbf{x}_i\} \in \mathcal{R}^{m \times n}$ denote a set of \mathbf{x}_i , where $i = 1, 2, 3, \dots, m$ and m is the number of patches extracted from the image.

Generally, the Euclidean distance is chosen as a similarity criterion. Given two image patches \mathbf{x}_i and \mathbf{x}_k , the similarity of them is formulated as

$$\|\mathbf{x}_i - \mathbf{x}_k\|_2^2 = \sum_{p=1}^n [x_{i(p)} - x_{k(p)}]^2 \quad (4)$$

Hence, using the Euclidean distance, the computational complexity of evaluating the similarity for any two patches of $\{\mathbf{x}_i\}$ is $\mathcal{O}(mn)$ that is fairly high when the value of \sqrt{n} is

6 or 8. So we try to find a feasible way called PCA-subspace Euclidean distance to measure the similarity in a lower domain.

Furthermore, the Euclidean distance is so simple that it fails to capture the intrinsic structure of image patches. For this drawback, we replace the Euclidean distance with SSIM distance to make full use of image structure information.

SSIM index as a similarity criterion

The idea of SSIM index can be traced back by Wang et al.²⁷ In that article, the author pointed out that natural image signals are highly structured, that is, their pixels exhibit strong dependencies, especially when they are spatially proximate, and these dependencies carry important information about the structure of the objects in the visual scene. And then SSIM is developed as a well-known quality metric used to measure the similarity between two images.

Given a reference image $\mathbf{x} \in \mathcal{R}^N$ and a test image $\mathbf{y} \in \mathcal{R}^N$, the SSIM between them is defined as

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [\ell(\mathbf{x}, \mathbf{y})]^\alpha \cdot [c(\mathbf{x}, \mathbf{y})]^\beta \cdot [s(\mathbf{x}, \mathbf{y})]^\gamma \quad (5)$$

where

$$\begin{aligned} \ell(\mathbf{x}, \mathbf{y}) &= \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \\ c(\mathbf{x}, \mathbf{y}) &= \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \\ s(\mathbf{x}, \mathbf{y}) &= \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \end{aligned} \quad (6)$$

The first formula in equation (6) is a function $\ell(\mathbf{x}, \mathbf{y})$ which gives luminance comparison. The second and the third formulas in equation (6) are two functions $c(\mathbf{x}, \mathbf{y})$ and $s(\mathbf{x}, \mathbf{y})$ which give contrast comparison and structure comparison, respectively. $\alpha > 0$, $\beta > 0$, and $\gamma > 0$ in equation (5) are parameters to acquire a proper weightage of the three functions. μ_x and μ_y are the mean luminance of images \mathbf{x} and \mathbf{y} , respectively. σ_x and σ_y stand for the standard deviations of the pixel distribution of images \mathbf{x} and \mathbf{y} , respectively. The covariance between \mathbf{x} and \mathbf{y} is denoted by σ_{xy} . And the positive constants C_1 , C_2 , and C_3 are used to avoid a null denominator. In general, we set $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$ to obtain a simplified form of the SSIM

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (7)$$

The value of SSIM ranges from 0 to 1, and the closer that the value is to 1, the higher the similarity between \mathbf{x} and \mathbf{y} is.

To reveal SSIMs merit that it can assess perceptual image quality better by making full use of structural

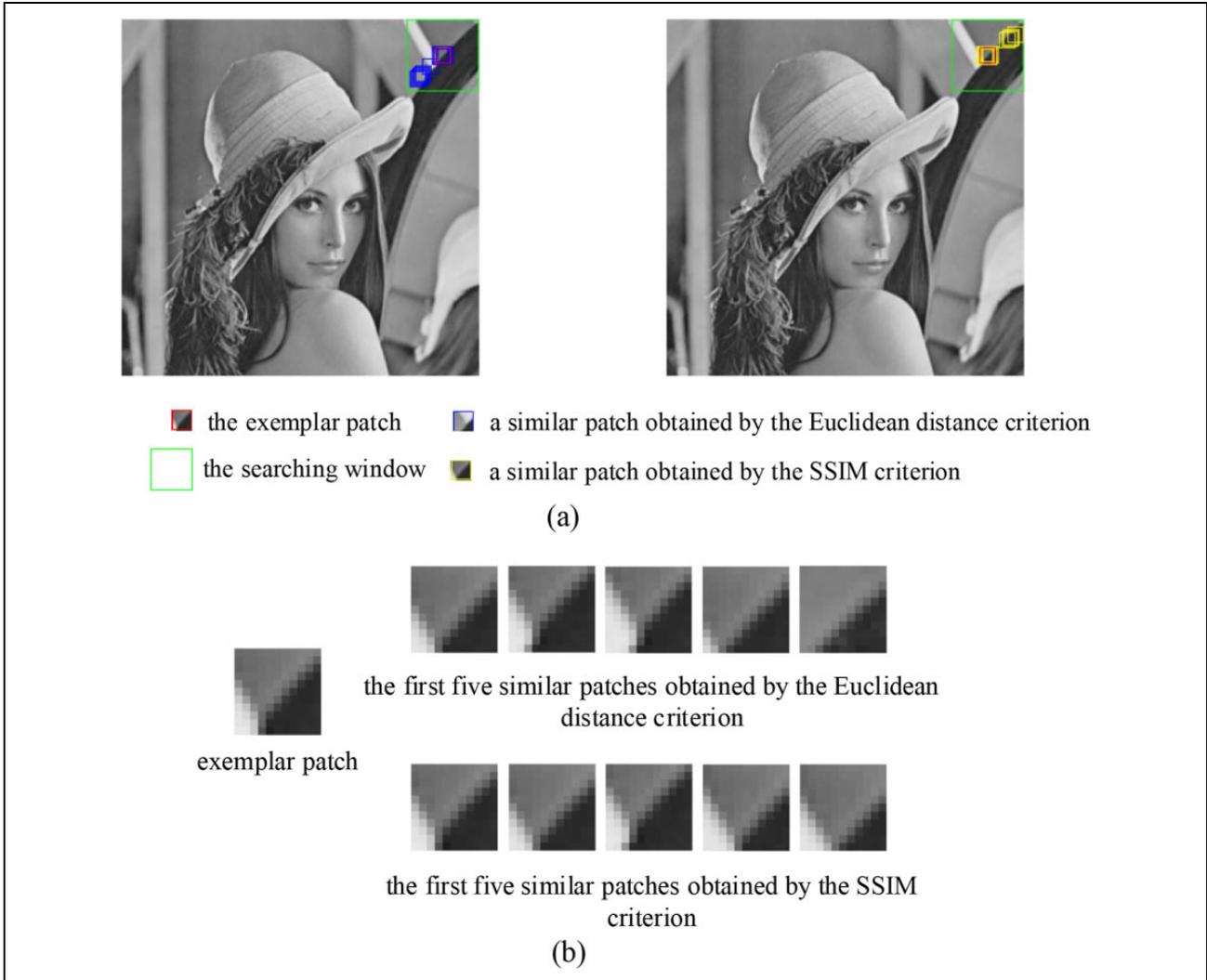


Figure 1. Comparison of searching similar image patches with the Euclidean distance-based and the SSIM distance-based criteria for the same exemplar patch. (a) Left: The first 18 similar patches searched in a search window by the Euclidean distance criteria. Right: The first 18 similar patches searched in a search window by the SSIM distance criteria. (b) Top: the first five similar patches extracted from the patches (in the searching window of the left image of (a)). Bottom: the first five similar patches extracted from the patches (in the searching window of the right image of (a)). SSIM: structural similarity.

information of images than the Euclidean distance, a motivating example is given by Wang et al.,²⁷ where the original “Boat” image is altered with different distortions which can be seen to have drastically different perceptual quality. Surprisingly, the distorted images yield nearly identical mean square error (MSE) relative to the original image. According to the same MSE value, the distorted images with drastically different perceptual quality tend to be considered with the same quality level. This unreasonable phenomenon further illustrates the defect of using the Euclidean distance to measure the similarity. By contrast, the SSIM values of the four distorted images are different from each other and can assess perceptual image quality.

This example inspires us to extend SSIM as a similarity measure to search similar image patches in an image.

Consequently, for two image patches $x_i \in \mathcal{R}^n$ and $x_k \in \mathcal{R}^n$, the SSIM distance of them is expressed as

$$\text{dis}_{\text{SSIM}}(x_i, x_k) = \frac{(2\mu_{x_i}\mu_{x_k} + C_1)(2\sigma_{x_i x_k} + C_2)}{(\mu_{x_i}^2 + \mu_{x_k}^2 + C_1)(\sigma_{x_i}^2 + \sigma_{x_k}^2 + C_2)} \quad (8)$$

The comparison of the Euclidean distance-based and the SSIM distance-based criteria for searching similar image patches to a given exemplar patch is shown in Figure 1. As shown in the left subfigure of Figure 1(a), the first 18 similar patches (small blue squares) searched by the Euclidean distance-based criterion are mainly distributed in the lower left of the exemplar patch (small red square). However, the first 18 similar patches (small yellow squares) searched by the SSIM distance-based criterion are more likely to be distributed in the upper right of the exemplar

patch (small red square), as shown in the right subfigure of Figure 1(a). Figure 1(b) represents the first five similar patches extracted from the 18 similar patches showed in the left and right subfigures of Figure 1(a), respectively. We can observe that based on the Euclidean distance metric, a good few configurations of the patch intensity on the corresponding positions of each patch are mismatched. In contrast, with the SSIM distance metric, most of the configurations of the patch intensities on the corresponding positions are the same or similar. Therefore, the patch searching method based on the SSIM distance can differentiate the patches more precisely than the one based on the Euclidean distance.

PCA-subspace Euclidean distance

What inspires us to propose this criterion is the assumption that image patch vectors focus on a lower dimensional manifold rather than the full space. This assumption stems from the researches done by Huang and Mumford³³ and Lee et al.³⁴ who found the distribution of data is extremely “sparse” with the majority of data points concentrated on clusters and nonlinear low-dimensional manifolds.

Meanwhile, PCA has been widely used in numerous image processing applications. When applied in image denoising,³⁵ PCA is computed as a proper local basis set for acquiring the noisy signal projection on it, and denoising is achieved by safely setting the small high-frequency coefficients to zero. This further encourages us to think whether PCA can be used to dramatically reduce the dimensionality of image patch vectors before evaluating their similarity. Based on the above consideration, some works^{36,37} have been done.

With all the above worthy work done by predecessor, we propose to replace the Euclidean distance as a similarity metric defined in equation (4) with a PCA-subspace Euclidean distance computed from projections of \mathbf{x}_i onto a low-dimensional subspace determined by PCA. The details of PCA-subspace Euclidean distance will be introduced in the following.

Let a matrix $\mathbf{X} \in \mathcal{R}^{m \times n}$ describe $\{\mathbf{x}_i\} \in \mathcal{R}^{m \times n}$ as follows

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mj} & \dots & x_{mn} \end{bmatrix} \quad (9)$$

where each row of \mathbf{X} corresponds to an image patch \mathbf{x}_i in $\{\mathbf{x}_i\}$, each column stands for a specific dimension, m is the number of image patches, and n is the dimensionality of image patch vectors.

First, we calculate the mean value of dimensions one by one and get a vector $\bar{\mathbf{d}} = [\bar{d}_1, \bar{d}_2, \dots, \bar{d}_j, \dots, \bar{d}_n]$, where \bar{d}_j is the mean value of j th dimension. Then, we subtract the mean value of dimensions from each corresponding column of \mathbf{X} and \mathbf{X} updates to

$$\bar{\mathbf{X}} = \begin{bmatrix} x_{11} - \bar{d}_1 & x_{12} - \bar{d}_2 & \dots & x_{1j} - \bar{d}_j & \dots & x_{1n} - \bar{d}_n \\ x_{21} - \bar{d}_1 & x_{22} - \bar{d}_2 & \dots & x_{2j} - \bar{d}_j & \dots & x_{2n} - \bar{d}_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i1} - \bar{d}_1 & x_{i2} - \bar{d}_2 & \dots & x_{ij} - \bar{d}_j & \dots & x_{in} - \bar{d}_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{m1} - \bar{d}_1 & x_{m2} - \bar{d}_2 & \dots & x_{mj} - \bar{d}_j & \dots & x_{mn} - \bar{d}_n \end{bmatrix} \quad (10)$$

Next, we can estimate its covariance matrix $\mathbf{W} \in \mathcal{R}^{n \times n}$ as

$$\mathbf{W} = \frac{1}{m-1} \bar{\mathbf{X}}^T \bar{\mathbf{X}} \quad (11)$$

Last, we compute the singular value decomposition³⁸ of \mathbf{W} as

$$\mathbf{W} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ u_{i1} & u_{i2} & \dots & u_{in} \\ \vdots & \vdots & \vdots & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \sigma_{n-1} & \vdots \\ 0 & 0 & \dots & \sigma_n \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{i1} & v_{i2} & \dots & v_{in} \\ \vdots & \vdots & \vdots & \vdots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{bmatrix} \quad (12)$$

where the matrix $\mathbf{\Sigma} \in \mathcal{R}^{n \times n}$ is a diagonal matrix with diagonal elements $\{\sigma_1, \sigma_2, \dots, \sigma_n\}$ called singular values, and the values $\{\sigma_k\}_{k=1}^n$ are always in decreasing order, that is, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{n-1} \geq \sigma_n \geq 0$. The matrix $\mathbf{U} \in \mathcal{R}^{n \times n}$ can be another simple form $\mathbf{U} = [u^{(1)}, u^{(2)}, \dots, u^{(n)}]$ and the

vectors $\{u^{(k)}\}_{k=1}^n$ consist of all the eigenvectors of the covariance matrix \mathbf{W} .

We can find a proper threshold δ to set the singular values $\{\sigma_k\}_{k=d+1}^n$ which satisfy $\sigma_k < \delta$ to zero, thus we get the low-dimensional PCA subspace

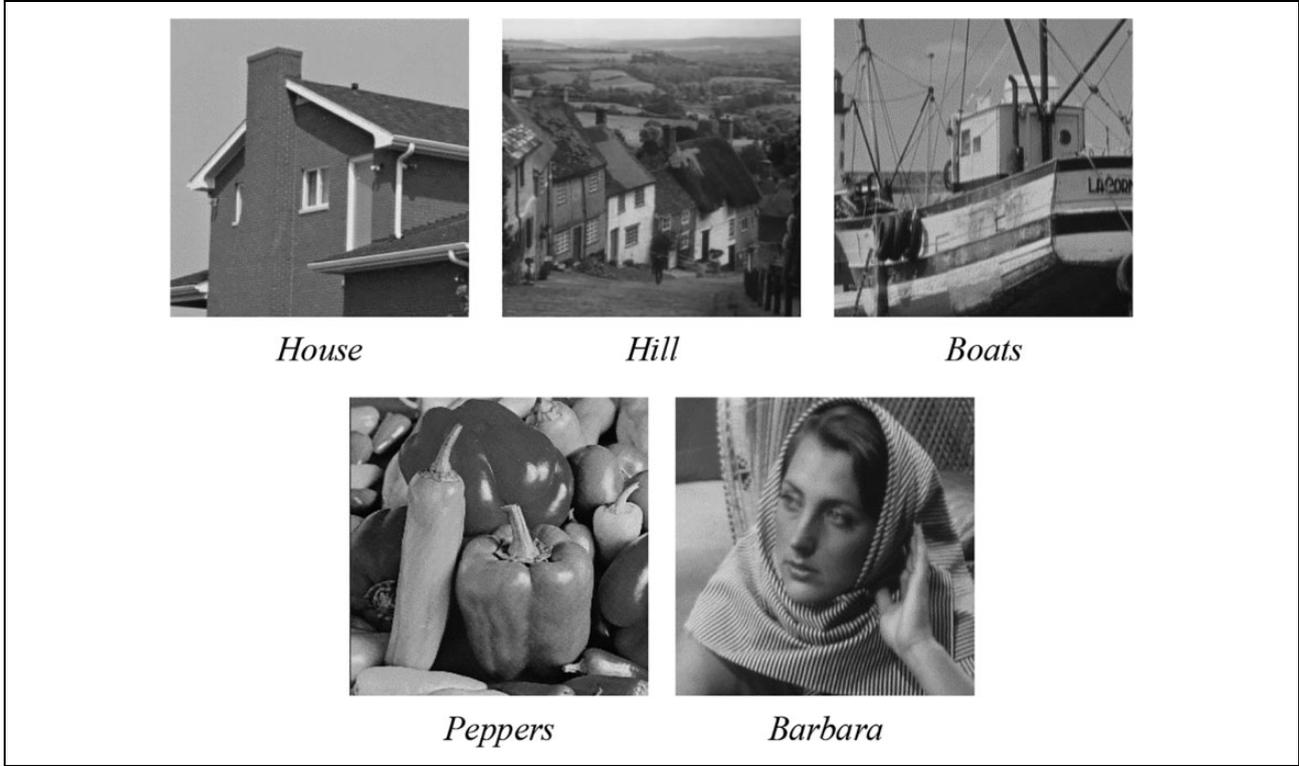


Figure 2. The original images used in Figure 3.

decided by $\mathbf{U}^{(d)} = [u^{(1)}, u^{(2)}, \dots, u^{(d)}]$ selected from $\mathbf{U} = [u^{(1)}, u^{(2)}, \dots, u^{(n)}]$, $0 < d \leq n$. And the projection of the image patch matrix \mathbf{X} onto this PCA subspace is given by

$$\mathbf{X}^d = \mathbf{X}\mathbf{U}^{(d)} = \begin{bmatrix} x_{1(1)}^d & x_{1(2)}^d & \dots & x_{1(j)}^d & \dots & x_{1(d)}^d \\ x_{2(1)}^d & x_{2(2)}^d & \dots & x_{2(j)}^d & \dots & x_{2(d)}^d \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{i(1)}^d & x_{i(2)}^d & \dots & x_{i(j)}^d & \dots & x_{i(d)}^d \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{m(1)}^d & x_{m(2)}^d & \dots & x_{m(j)}^d & \dots & x_{m(d)}^d \end{bmatrix} \quad (13)$$

Now, given two image patches \mathbf{x}_i and \mathbf{x}_k , the PCA-subspace Euclidean distance as a similarity metric can be defined as

$$\text{dis}_{\text{PCA}}(\mathbf{x}_i, \mathbf{x}_k) = \|\mathbf{x}_i^d - \mathbf{x}_k^d\|_2^2 = \sum_{p=1}^d [x_{i(p)}^d - x_{k(p)}^d]^2 \quad (14)$$

Figure 3 shows the top six eigenvectors of $\mathbf{U} = [u^{(1)}, u^{(2)}, \dots, u^{(36)}]$ computed from 6×6 image patches for the images “House,” “Hill,” “Boats,” “Peppers,” and “Barbara” (see Figure 2). The first eigenvector (the left-most column in Figure 3) corresponding to

the largest eigenvalue σ_1 is usually approximately flat. This flat eigenvector represents the average intensity in the 6×6 patch. The next two eigenvectors (columns 2–3 in Figure 3) almost always represent two orthogonal gradient directions which are necessary for representing edges. Generally, the front few eigenvectors of \mathbf{U} (columns 2–6 in Figure 3) represent edge and ridge patterns; however, in the case of strongly texture images, they can also represent the dominant texture patterns. For example, the image “Barbara” in Figure 2 owns abundant stripe patterns, so its 3th–6th eigenvectors (columns 3–6 in the last row in Figure 3) show specific texture of the image. Based on the above analysis that the front few PCA eigenvectors represent some edges and texture information of images, our experiments for deblurring conducted in “Experimental results” section demonstrate that the similarity evaluation in a low-dimensional subspace determined by PCA can capture the intrinsic structure of image patches more precisely than that in a full space.

Image deblurring via our proposed novel similarity criteria under the NCSR framework

The deblurring model under the NCSR framework

For an image \mathbf{X}_{im} of size $\sqrt{N} \times \sqrt{N}$ pixels and an image patch \mathbf{x}_{im_i} of size $\sqrt{n} \times \sqrt{n}$ pixels extracted from \mathbf{X}_{im} at location i , using an extraction matrix \mathbf{E}_i , that is,

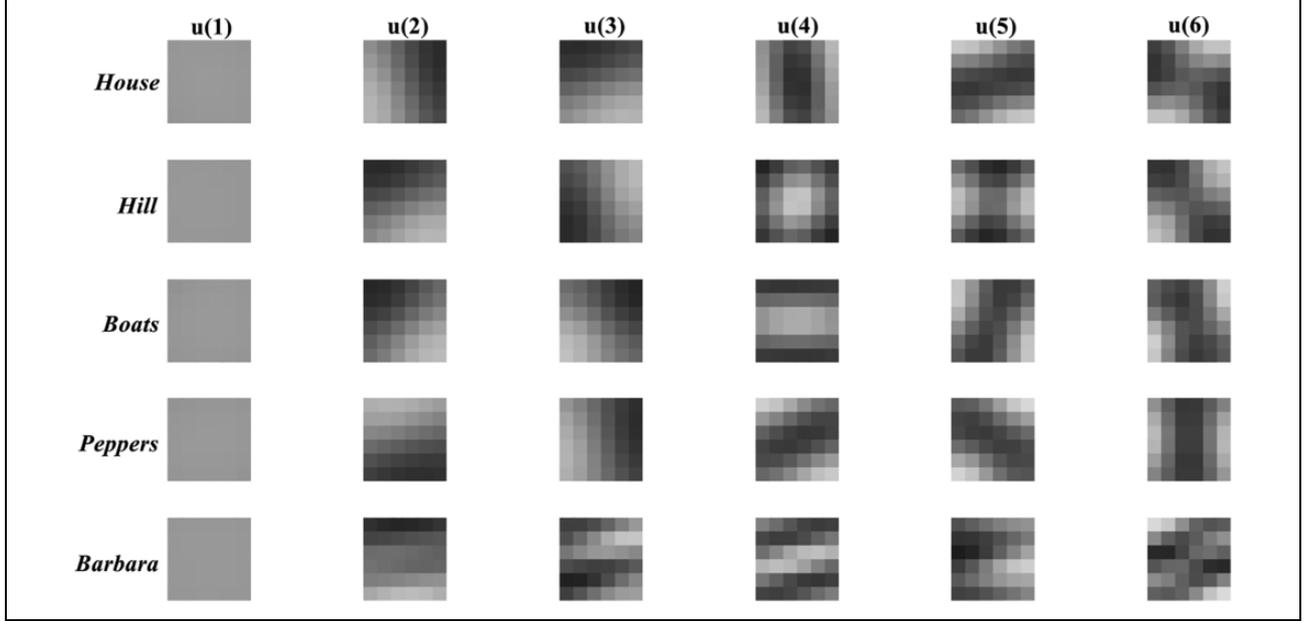


Figure 3. The top six eigenvectors for 6×6 image patches.

$\mathbf{x}_{im_i} = \mathbf{E}_i \mathbf{X}_{im}$. Let $\mathbf{x} = [x_{(1)}, x_{(2)}, \dots, x_{(N)}]^T \in \mathcal{R}^N$, $\mathbf{x}_i = [x_{i(1)}, x_{i(2)}, \dots, x_{i(n)}]^T \in \mathcal{R}^n$ denotes column vectors obtained by stacking the columns of \mathbf{X}_{im} and \mathbf{x}_{im_i} , respectively. According to equation (3), given a learned dictionary $\mathbf{D} \in \mathcal{R}^{n \times M}$ ($n \leq M$, M is the number of contained atoms in \mathbf{D}), each patch can be represented sparsely as $\mathbf{x}_i \approx \mathbf{D} \alpha_{x_i}$ by solving the l_1 -minimization-based regularization form

$$\alpha_{x_i} = \arg \min_{\alpha_i} \{ \|\mathbf{x}_i - \mathbf{D} \alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \} \quad (15)$$

Then, the entire image \mathbf{x} can be reconstructed through a straightforward least-square solution

$$\mathbf{x} \approx \mathbf{D} \circ \alpha_x = \left(\sum_{i=1}^N \mathbf{E}_i^T \mathbf{E}_i \right)^{-1} \sum_{i=1}^N (\mathbf{E}_i^T \mathbf{D} \alpha_{x_i}) \quad (16)$$

where α_x denotes the concatenation of all $\alpha_{x_i} \in \mathcal{R}^M$ and the notation \circ is used to simplify the operation of $\left(\sum_{i=1}^N \mathbf{E}_i^T \mathbf{E}_i \right)^{-1} \sum_{i=1}^N (\mathbf{E}_i^T \mathbf{D} \alpha_{x_i})$. Equation (16) shows that we can reconstruct the whole image \mathbf{x} by averaging each recovery patch of \mathbf{x}_i .

In the process of image deblurring, however, what we have is just the degraded observation $\mathbf{y} = \mathbf{H} \mathbf{x} + \mathbf{n}$. To

reconstruct the unknown clean image \mathbf{x} based on \mathbf{y} , we first code \mathbf{y} sparsely, that is, $\mathbf{y} \approx \mathbf{D} \circ \alpha_y$, and the sparse coding coefficient α_y is cast as

$$\alpha_y = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \mathbf{H} \mathbf{D} \circ \alpha\|_2^2 + \lambda \sum_i \|\alpha_i\|_1 \right\} \quad (17)$$

Then the image \mathbf{x} is reconstructed as $\mathbf{x} \approx \mathbf{D} \circ \alpha_y$.

It's obvious that if we want to obtain a higher quality image reconstruction, we must make the sparse coding coefficients α_y as close to α_x as possible. Consequently, another constraint to suppress the deviation between α_y and α_x is needed urgently. Considering this, a concept of SCN is proposed and defined as

$$\mathbf{V}_\alpha = \alpha_y - \alpha_x \quad (18)$$

Some experiments performed by Dong et al.²⁰ show that the SCN \mathbf{V}_α can be well characterized by Laplacian functions, which motivate us to model \mathbf{V}_α with a Laplacian prior.

Now, another problem is that α_x is unavailable. Nonetheless, we can find a way to acquire a good estimation of α_x , denoted by $\hat{\alpha}_x$. Then, $\hat{\mathbf{V}}_\alpha = \alpha_y - \hat{\alpha}_x$ can be a good estimation of \mathbf{V}_α . Combining \mathbf{V}_α as a constraint to equation (17), the more accurate sparse coding model is

$$\alpha_y = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \mathbf{H} \mathbf{D} \circ \alpha\|_2^2 + \lambda \sum_i \|\alpha_i\|_1 + \gamma \sum_i \|\alpha_i - \hat{\alpha}_{x_i}\|_1 \right\} \quad (19)$$

In addition, in this article, considering that the contents can vary significantly across different images or different patches in a single image, we adopt an adaptive subdictionary selection strategy that learn a series of PCA subdictionaries from

pre-collected k clusters of image patches and then, for a given patch to be processed, select one subdictionary adaptively to characterize the local sparse domain. Consequently, the sparsity of α_y can be ensured and the local sparse regularization

term $\sum_i \|\alpha_i\|_1$ can be removed. The final nonlocal sparse regularization-based deblurring model is shown as

$$\alpha_y = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \mathbf{HD} \circ \alpha\|_2^2 + \gamma \sum_i \|\alpha_i - \hat{\alpha}_{x_i}\|_1 \right\} \quad (20)$$

$$\alpha_y = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \mathbf{HD} \circ \alpha\|_2^2 + 2\sqrt{2}\sigma_n^2 \times \sum_i \sum_j \frac{1}{\sigma_{i(j)}} |\alpha_{i(j)} - \beta_{i(j)}| \right\} \quad (21)$$

Compared equation (20) with equation (21), we have

$$\begin{aligned} \gamma_i &= [\gamma_{i(1)}, \gamma_{i(2)}, \dots, \gamma_{i(j)}, \dots, \gamma_{i(M)}] \\ &= \left[\frac{2\sqrt{2}\sigma_n^2}{\sigma_{i(1)}}, \frac{2\sqrt{2}\sigma_n^2}{\sigma_{i(2)}}, \dots, \frac{2\sqrt{2}\sigma_n^2}{\sigma_{i(j)}}, \dots, \frac{2\sqrt{2}\sigma_n^2}{\sigma_{i(M)}} \right] \end{aligned} \quad (22)$$

The deblurring model solving via our novel similarity criteria

The deblurring model can be solved in an iterative way, and the main procedures can be generalized in algorithm 1, which is corresponding to Figure 4.

Experimental results

In this section, extensive experimental results are showed in detail to verify the performance of nonlocal sparse regularization-based image deblurring via novel similarity criteria. The parameter setting is as follows: the image patch size is $\sqrt{n} \times \sqrt{n} = 6 \times 6$, the width of overlapping between adjacent patches is 2 pixels, the number of clusters is $K = 64$, the number of similar patches best matched a given patch in a searching window is set to be $Q = 13$, the threshold that defined in ‘‘PCA-subspace Euclidean distance’’ section is set to be $\delta = 100$ to get the low-dimensional PCA subspace, and the outer and inner loop numbers are $L = 6$ and $J = 120$, respectively. Besides, the regularization parameter γ can be automatically determined by equation (22). All the experiments are conducted in MATLAB 2016b on a PC with Intel(R) Core(TM) i7-4790 CPU processor (3.60 GHz), 32.00G memory, and Windows 7 operating system.

To evaluate the quality of deblurred images, in addition to the PSNR generally used to evaluate the objective image quality, two more powerful perceptual quality metrics SSIM and FSIM are adopted to evaluate the visual quality. For color images, image deblurring operations are only applied to the luminance component. All the experimental test images are listed in Figure 5. They are selected from an image database (available at <http://decsai.ugr.es/cvg/dbi-magenes/index.php>). Due to the limit of space, only parts

Let’s define $\phi_i = \alpha_{y_i} - \hat{\alpha}_{x_i}$, $\phi_{i(j)}$ is the j^{th} element of ϕ_i , and $\sigma_{i(j)}$ is the standard deviation of $\phi_{i(j)}$. In a maximum a posteriori estimator perspective (refer to the study by Zhang et al.²⁵ for more details), the sparse codes α_y can be represented as

of the experimental results are shown in this article. Our MATLAB code can be downloaded at the website: https://github.com/wangnannanying/INSR_Deblur-SR.

Deblurring experiments for the simulated blurred images

In this subsection, we conduct two sets of experiments to demonstrate the performance of our proposed method for image deblurring. In the first set, two commonly used blur kernels, that is, a 9×9 uniform kernel and a two-dimensional Gaussian blur kernel with standard deviation of 1.6 are exploited for simulating. Then, the blurred images are further corrupted by additive Gaussian noise with standard deviation of $\sigma_n = \sqrt{2}$. In the second set, we consider six typical deblurring scenarios used as the benchmarks in many publications.^{39,40} The blur PSF and the variance of the noise σ_n^2 for each scenario are summarized in Table 1. Each of the scenarios was tested with the four standard images, that is, Cameraman, House, Lena, and Barbara.

We compare our deblurring method with four state-of-the-art deblurring methods, including fast iterative shrinkage/thresholding algorithm (FISTA),⁴¹ iterative decoupled deblurring block-matching and 3D filtering algorithm (IDD-BM3D),⁴⁰ NCSR,²⁰ and group-based sparse representation (GSR).²¹ FISTA introduced new gradient-based schemes for the constrained TV-based image deblurring. IDD-BM3D is an improved version of BM3D.⁴² NCSR proposed a centralized sparse constraint, which exploits the image nonlocal redundancy to reduce the SCN.²⁰ GSR sparsely represented natural images in the domain of group and characterized the intrinsic local sparsity and nonlocal self-similarity of natural images simultaneously in a unified manner.

The subjective image visual comparisons of our proposed method and other four deblurring methods on images ‘‘House’’ and ‘‘Starfish’’ are shown in Figures 6 and 7, from which several conclusions can be drawn: (1) FISTA fails to suppress the noise and there exists noticeable ringing effects surround strong edges (see the House contour in Figure 6(b) and the latticed texture in Starfish’s surface

Algorithm 1: An improved nonlocal sparse regularization-based image deblurring**Input**

Input the observed image y and the degraded operator \mathbf{H} .

Initialization

Set the initial estimate as $\hat{x} = y$;

Set the initial regularization parameter γ ;

Outer loop (iterate on $\ell = 1, 2, \dots, L$)

(a) Extracting image patches

With a step size 1, divide the blurred image x into m' overlapped patches $x_i' \in \mathcal{R}^n$, $i = 1, 2, \dots, m'$. Using these patches, we obtain a PCA dictionary $\mathbf{D}_0 \in \mathcal{R}^{n \times d}$ which equals to $\mathbf{U}^{(d)}$ in equation (13).

(b) Clustering

For each image patch x_i' , obtain its projection onto PCA subspace by computing $x_i'^{\ell} = \mathbf{D}_0 x_i'$. Then, choose K image patches randomly in PCA subspace as the initial mean of the clusters, i.e. $\bar{x}_{c'}'$, $c = 1, 2, \dots, K$.

Cluster the m' overlapped patches into K clusters by using the k -means clustering method. More exactly, for a given patch x_i' , calculate the PCA-subspace Euclidean distance between x_i' and each mean of the clusters $\bar{x}_{c'}'$, $c = 1, 2, \dots, K$, based on equation (14). Then, the cluster that x_i' falls into is the one with the minimum of the PCA-subspace Euclidean distances.

(c) Learning the PCA dictionary

For each cluster acquired in Step 2, learn a dictionary $\mathbf{D} = \{\mathbf{D}_k\}_{k=1}^K$ of PCA bases called PCA subdictionaries. Since the patches in a cluster is similar to each other, so we learn a compact PCA dictionary rather than an over-complete dictionary.

(d) Select the PCA subdictionary for each patch

For each patch x_i in $m(m = m'/4)$ overlapped patches extracted from the image x with a step size 2, select a PCA subdictionary learned in Step 3 to code it. The way to subdictionary selection is first computing the SSIM distance between x_i and each mean of the clusters \bar{x}_c , $c = 1, 2, \dots, K$, based on equation (8), and then choosing the PCA subdictionary learned from the cluster whose mean corresponds to the smallest SSIM distance.

(e) Inner loop (iterate on $j = 1, 2, \dots, J$)**(1) Block matching**

Search the Q nonlocal similar patches to each image patch x_i in a large $S \times S$ window centered at location i to obtain the similar patches set $\{x_{i,q}\}_{q=0}^{Q-1}$ which includes the image patch x_i itself, i.e., $x_i = x_{i,0}$. The search method for similar patches in $\{x_{i,q}\}_{q=0}^{Q-1}$ is similar to the way used in Step 2 for clustering. To be specific, choose Q patches whose PCA-subspace Euclidean distance to x_i is the first- Q smallest compared with all patches in the searching window.

(2) Sparse coding

Use the selected subdictionaries to code all patches in $\{x_{i,q}\}_{q=0}^{Q-1}$ and the corresponding sparse coding coefficients can be calculated as:

$$\alpha_{i,q} = \mathbf{D}_{i,q}^T x_{i,q}^T, \alpha_{i,q} \in \{\alpha_{x_{i,0}}\}_{q=0}^{Q-1} \quad (23)$$

(3) Estimation for sparse coding coefficients $\{\alpha_{x_i}\}$

Similar to the nonlocal means approach in image denoising, we estimate α_{x_i} as the weighted average of $\alpha_{i,q}$:

$$\hat{\alpha}_{x_i} = \sum_{\alpha_{i,q} \in \{\alpha_{i,q}\}} w_{i,q} \alpha_{i,q} \quad (24)$$

and set the weights to be inversely proportional to the distance between patches x_i and $x_{i,q}$:

$$w_{i,q} = \frac{1}{W} e^{-\frac{\|x_i - x_{i,q}\|_2^2}{h}} \quad (25)$$

where W is a normalization factor, h is a predetermined scalar to control the decay of the exponential expression.

For all the m overlapped patches, we can get a sparse coding coefficients estimation set $\{\hat{\alpha}_{x_i}\}$.

(4) Solve the model by iterative shrinkage algorithm

In the j th iteration, the shrinkage operator for the $\alpha_y^{(j)}$ is:

$$\alpha_y^{(j)} = \{\alpha_{y_i}\}^{(j)} = \left\{ S_{\tau} \left(y_{y_i}^{(j)} - \hat{\alpha}_{x_i}^{(j)} \right) + \hat{\alpha}_{x_i}^{(j)} \right\} \quad (26)$$

where $S_{\tau}(\cdot)$ is the classic soft-thresholding operator, $\tau = \gamma_i/c$ and $y^{(j)} = (\mathbf{HD})^T \left(y - (\mathbf{HD}) \circ \alpha_y^{(j-1)} \right) / c + \alpha_y^{(j-1)}$, where c is an auxiliary parameter to guarantee the surrogate function convergent. The interesting readers can refer to¹⁹ for more details about the shrinkage operator.

(5) Patch to image transformation

The entire image can be reconstructed by $x^{(j)} \approx \mathbf{D} \circ \alpha_y^{(j)}$.

End for

End for

Output

The reconstructed image $x^{(k)}$.

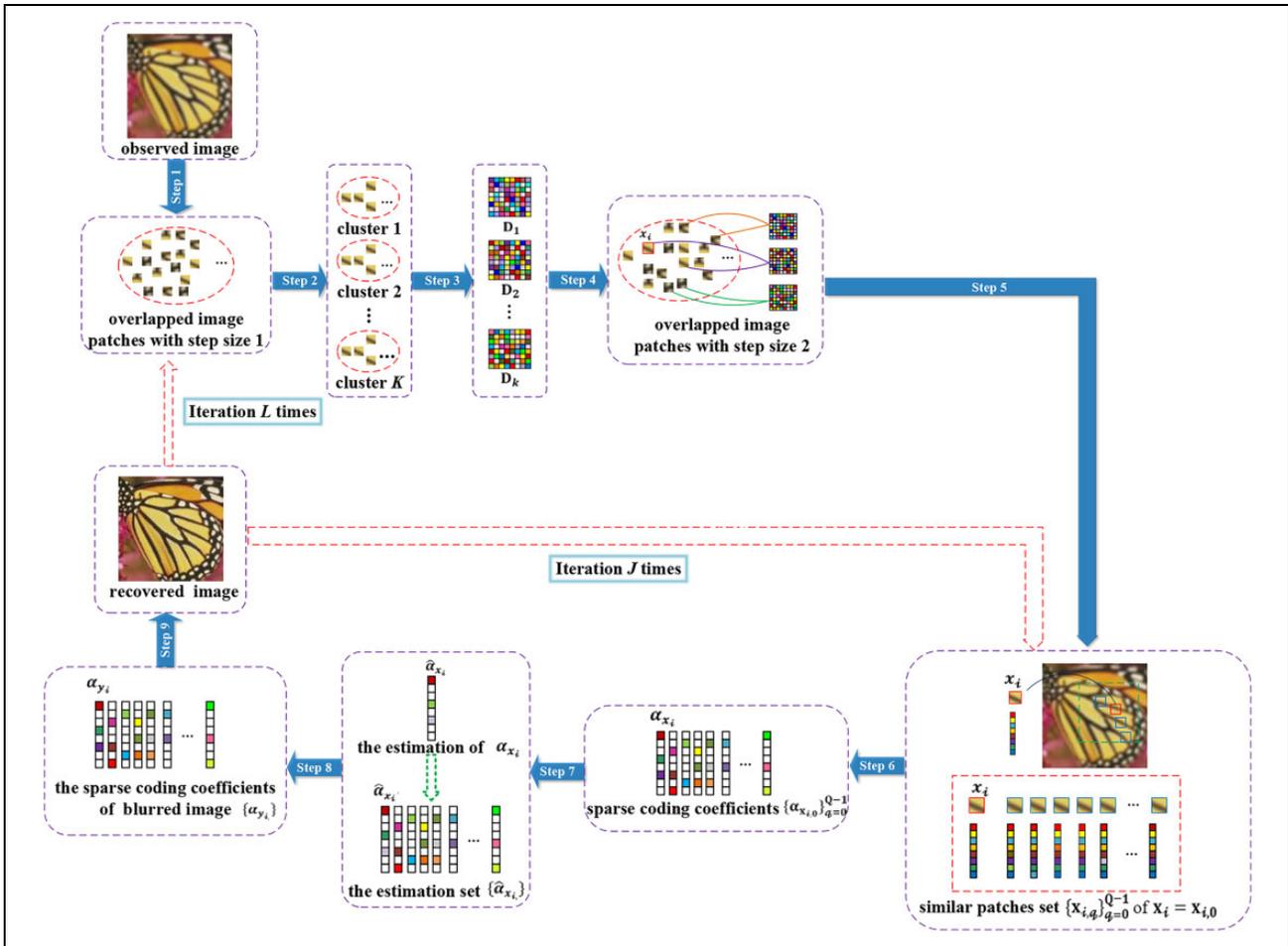


Figure 4. The flowchart of nonlocal sparse regularization-based image deblurring via novel similarity criteria.

in Figure 7(b)). The deblurring performance is unsatisfactory. (2) IDD-BM3D, NCSR, and GSR achieve deblurred images with similar quality (all noticeably better than FISTA), of which NCSR works quite well in clearly restoring large edges without any noticeable ringing artifacts and exhibits powerful ability in removing noise in smooth and low-activity regions. However, the price is paid in loss of some image details, leading to blurs in texture regions. Moreover, there still remains some noticeable noise around edges (see the contour of eaves in Figure 6(d) and the contour of latticed texture in Figure 7(d)). (3) As expected, our proposed method generates near-perfect deblurring results that most of the image edges and textures are restored very well, while the noise existing in image has been effectively suppressed. Compared to NCSR, our method recovers much cleaner and sharper image edges and textures (see Figures 6(f) and 7(f)). Such experimental findings clearly suggest that our proposed model is a stronger prior for the class of photographic images containing strong edges/textures.

The PSNR, SSIM, and FSIM comparison results on 10 test images (see Figure 5) in the first set of experiments among five competing methods are reported in Table 2 for

the uniform blur and Table 3 for the Gaussian blur, respectively. From Tables 2 and 3, it can be observed that our method (Ours) clearly outperforms all other four for most of 10 test images. The gains are mostly impressive for “Butterfly” and “Leaves” images which contain abundant strong edges or textures. One possible explanation is that our method is capable of striking a better trade-off between exploiting local and nonlocal dependencies within those images. The proposed method achieves superior performance to other competing methods, and our method outperforms FISTA by 2.56 and 2.13 dB for the uniform blur and Gaussian blur, respectively. It’s obvious that IDDBM3D, NCSR, and GSR produce very similar results and the performance measured by PSNR, SSIM, and FSIM is improved significantly, compared with FISTA. What is exciting is that our model outperforms all these competitive methods, and our model outperforms IDDBM3D, NCSR, and GSR by (0.71 dB, 0.43 dB), (0.41 dB, 0.26 dB), and (0.51 dB, 0.47 dB) for uniform blur and Gaussian blur, respectively, which is consistent with the subjective visual comparisons shown in Figures 6 and 7.

In Table 4, we present the comparison of improvement of signal-to-noise ratio (ISNR) values achieved by each



Figure 5. All experimental test images (256×256).

Table 1. Various blur PSFs and noise variances used in six typical non-blind deblurring experiments in the second set.

Scenario	PSF	σ_n^2
1	$1/(1+z_1^2+z_2^2), z_1, z_2 = -7, \dots, 7$	2
2	$1/(1+z_1^2+z_2^2), z_1, z_2 = -7, \dots, 7$	8
3	$[1\ 4\ 6\ 4\ 1]^T [1\ 4\ 6\ 4\ 1]/256$	49
4	9×9 uniform	≈ 0.3
5	Gaussian with std = 1.6	4
6	Gaussian with std = 0.4	64

PSF: point spread function.

deblurring method for four test images. The ISNR of the image is another common measurement in image restoration and is defined as

$$\text{ISNR} = 10 \log_{10} \frac{\sum_{i=1}^N (\mathbf{x}^{(i)} - \mathbf{y}^{(i)})^2}{\sum_{i=1}^N (\mathbf{x}^{(i)} - \hat{\mathbf{x}}^{(i)})^2} \quad (27)$$

From Table 4, it is clear to observe that our method achieves the highest ISNR results in all the six scenarios

when deblurring the blurred images “Cameraman” and “Lena” and obtains the highest ISNR values in the four scenarios when deblurring the blurred image House. However, the inferior ISNR results appear when deblurring the blurred image “Barbara,” which we guess is because our method has a possibly high sensitivity to the type of PSF for images with rich textures, which is needed to be ameliorated in our future works.

Ablation studies of SSIM distance and PCA-subspace Euclidean distance

In this article, our method adopts both SSIM distance and PCA-subspace Euclidean distance to attain better performance of deblurring. In order to study further the effects of these two similarity criteria on the deblurring results separately, two experiments, of which one is deblurring only with SSIM distance (called Ours-SSIM) and another is deblurring only with PCA-subspace Euclidean distance (called Ours-PCA), were conducted and compared to our method (see Table 5). As described in “Image deblurring via our proposed novel similarity criteria under the NCSR framework” section, in this article, we merely conduct image deblurring via our proposed similarity criteria under

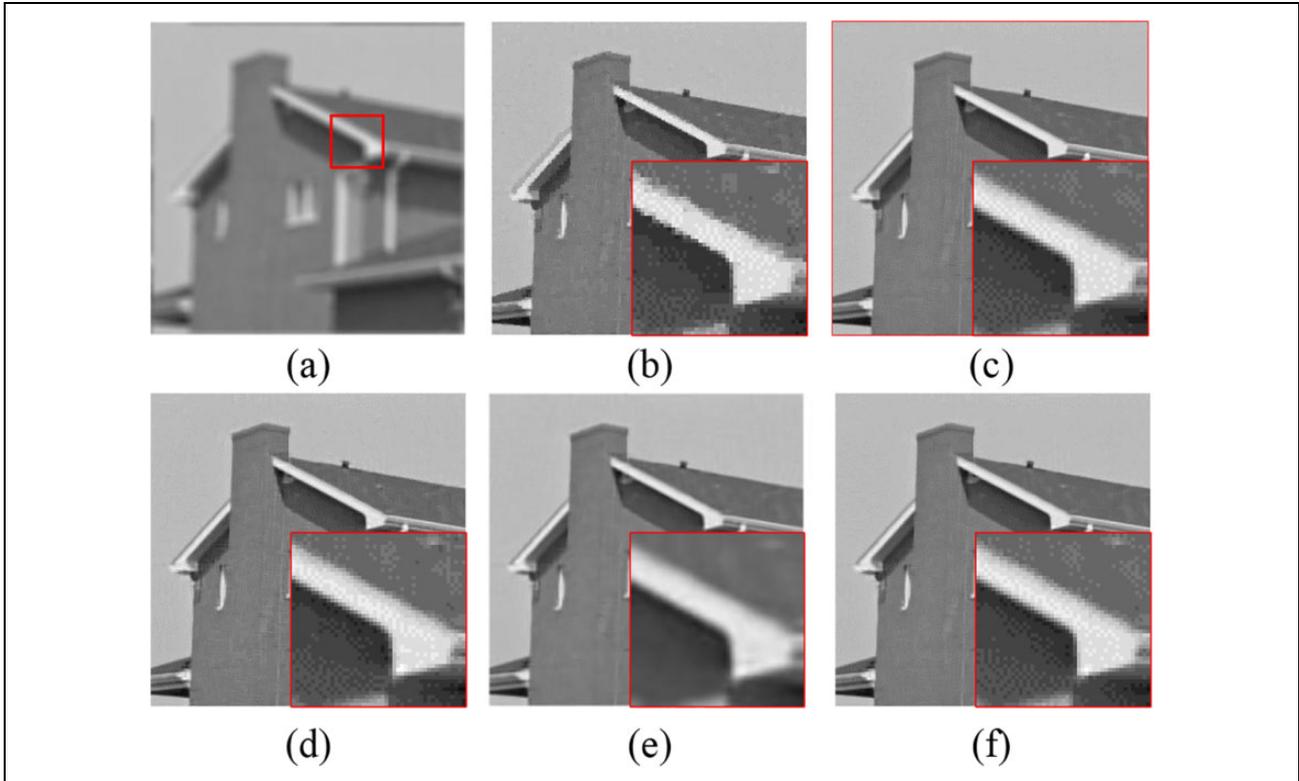


Figure 6. Visual quality comparison of image deblurring on the gray image House (256×256). (a) The noisy and blurred image (9×9 uniform blur, $\sigma_n = \sqrt{2}$). (b) The deblurred image by FISTA (PSNR = 31.99 dB; SSIM = 0.8490; FSIM = 0.9017). (c) The deblurred image by IDD-BM3D (PSNR = 34.44 dB; SSIM = 0.8786; FSIM = 0.9369). (d) The deblurred image by NCSR (PSNR = 34.31 dB; SSIM = 0.8755; FSIM = 0.9415). (e) The deblurred image by GSR (PSNR = 34.48 dB; SSIM = 0.8782; FSIM = 0.9403). (f) The deblurred image by our method (PSNR = 34.61 dB; SSIM = 0.8806; FSIM = 0.9369). FISTA: fast iterative shrinkage/thresholding algorithm; PSNR: peak signal-to-noise; SSIM: structural similarity; FSIM: feature similarity; IDD-BM3D: iterative decoupled deblurring block-matching and 3D; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

the superior NCSR framework. To highlight the superior performance of our method, we also list the result of NCSR in Table 5. Besides, we simply show the PSNR values for a concise comparison in Table 5. From Table 5, we can conclude that whether using only SSIM distance (Ours-SSIM) or using only PCA-subspace Euclidean distance (Ours-PCA), better performance than NCSR can be achieved, while even better results can be achieved by using both similarity criteria (Ours). This indicates that SSIM distance or PCA-subspace Euclidean distance we have proposed in this article is solely for improving the deblurring performance. In addition, the result using both similarity criteria (Ours) acquired the superior performance shows that the two similarity criteria can promote each other.

Algorithm stability

Here, we provide empirical evidence to illustrate the stability of the proposed method. Take the cases of image deblurring for two blur types in the first sets of experiments as examples. Figure 8(a) and (b) plots the evolutions of PSNR versus outer loop iteration numbers for five test images in the cases of image deblurring for uniform blur and Gaussian blur, respectively.

It is observed that with the growth of outer loop iteration numbers, all the PSNR curves of five test images increase monotonically when the outer loop iteration numbers increase from one to six and ultimately become flat and stable when the outer loop iteration numbers are greater than six for both the uniform blur and the Gaussian blur, exhibiting good stability of the proposed method. Based on the above results, the outer loop iteration numbers are set to be six in our experiments described in “Experimental results” section.

Effect of number of clusters and number of best matched patches

This subsection will give some discussion about the deblur performance affected by K and Q which are the number of clusters and the number of best matched patches, respectively.

To investigate the sensitivity of K and Q , experiments with reference to various K and Q in the case of image deblurring for three test images “Barbara,” “House,” “Parrot” are conducted. The performance comparisons with various K and Q in the case of image deblurring with 9×9

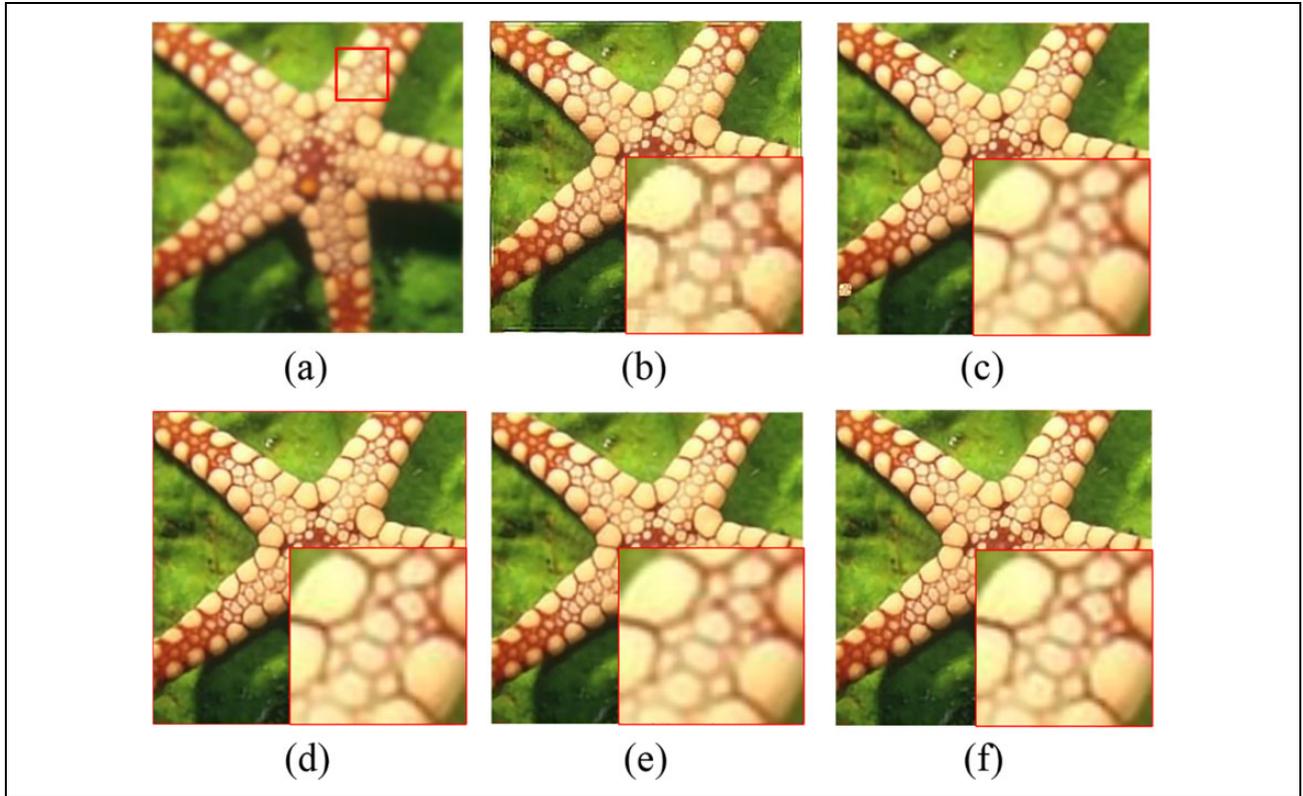


Figure 7. Visual quality comparison of image deblurring on the color image Starfish (256×256). (a) The noisy and blurred image (Gaussian blur, $\sigma_n = \sqrt{2}$). (b) The deblurred image by FISTA (PSNR = 29.42 dB; SSIM = 0.8349; FSIM = 0.9256). (c) The deblurred image by IDD-BM3D (PSNR = 31.66 dB; SSIM = 0.9156; FSIM = 0.9496). (d) The deblurred image by NCSR (PSNR = 32.27 dB; SSIM = 0.9229; FSIM = 0.9551). (e) The deblurred image by GSR (PSNR = 31.61 dB; SSIM = 0.9165; FSIM = 0.9471). (f) The deblurred image by our method (PSNR = 32.50 dB; SSIM = 0.9270; FSIM = 0.9564). FISTA: fast iterative shrinkage/thresholding algorithm; PSNR: peak signal-to-noise; SSIM: structural similarity; FSIM: feature similarity; IDD-BM3D: iterative decoupled deblurring block-matching and 3D; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

Table 2. PSNR (dB), SSIM, and FSIM comparisons by different deblurring methods for the uniform blur in the first set.

9×9 Uniform blur with Gaussian noise $\sigma_n = \sqrt{2}$												
Images		Butterfly	Boats	C. Man	House	Parrot	Lena	Barbara	Starfish	Peppers	Leave	Avg.
FISTA ⁴¹	PSNR	28.37	29.04	26.82	31.99	29.11	28.33	25.75	27.75	28.43	26.49	28.21
	SSIM	0.9058	0.8355	0.8278	0.8490	0.8750	0.8274	0.7440	0.8200	0.8134	0.9023	0.8400
	FSIM	0.9119	0.8858	0.8627	0.9017	0.9002	0.8798	0.8375	0.8775	0.8813	0.8958	0.8834
IDD-BM3D ⁴⁰	PSNR	29.21	31.20	28.56	34.44	31.06	29.70	27.98	29.48	29.62	29.38	30.06
	SSIM	0.9216	0.8820	0.8580	0.8786	0.9041	0.8654	0.8225	0.8640	0.8422	0.9418	0.8780
	FSIM	0.9287	0.9304	0.9007	0.9369	0.9364	0.9197	0.9014	0.9167	0.9200	0.9295	0.9220
NCSR ²⁰	PSNR	29.68	31.08	28.62	34.31	31.95	29.96	28.10	30.28	29.66	29.98	30.36
	SSIM	0.9273	0.8810	0.8574	0.8755	0.9103	0.8676	0.8255	0.8807	0.8402	0.9485	0.8814
	FSIM	0.9271	0.9294	0.9026	0.9415	0.9411	0.9254	0.9117	0.9293	0.9220	0.9341	0.9263
GSR ²¹	PSNR	28.94	31.34	28.28	34.48	31.60	30.10	28.95	29.90	29.66	29.36	30.26
	SSIM	0.9210	0.8860	0.8538	0.8782	0.9083	0.8771	0.8487	0.8744	0.8484	0.9424	0.8838
	FSIM	0.9151	0.9326	0.8937	0.9403	0.9418	0.9281	0.9227	0.9217	0.9231	0.9259	0.9245
Ours	PSNR	30.60	31.40	28.71	34.61	32.57	30.21	27.96	30.85	29.94	30.86	30.77
	SSIM	0.9363	0.8906	0.8646	0.8806	0.9157	0.8808	0.8270	0.8927	0.8527	0.9546	0.8896
	FSIM	0.9383	0.9328	0.9044	0.9369	0.9446	0.9267	0.9250	0.9355	0.9241	0.9432	0.9312

The bold values mean the largest ones. FISTA: fast iterative shrinkage/thresholding algorithm; PSNR: peak signal-to-noise; SSIM: structural similarity; FSIM: feature similarity; IDD-BM3D: iterative decoupled deblurring block-matching and 3D; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

Table 3. PSNR (dB), SSIM, and FSIM comparisons by different deblurring methods for the Gaussian blur in the first set.

Gaussian blur 1.6 with Gaussian noise $\sigma_n = \sqrt{2}$												
Images		Butterfly	Boats	C. Man	House	Parrot	Lena	Barbara	Starfish	Peppers	Leaves	Avg.
FISTA ⁴¹	PSNR	30.36	29.36	26.80	31.50	31.23	29.47	25.03	29.42	28.43	29.33	29.22
	SSIM	0.9374	0.8509	0.8241	0.8254	0.9066	0.8537	0.7377	0.8349	0.8134	0.9480	0.8532
	FSIM	0.9452	0.9024	0.8845	0.8968	0.9290	0.9011	0.8415	0.9256	0.9057	0.9393	0.9071
IDD-BM3D ⁴⁰	PSNR	30.73	31.68	28.17	34.08	32.89	31.45	27.19	31.66	29.99	31.40	30.92
	SSIM	0.9469	0.9036	0.8705	0.8820	0.9319	0.9103	0.8231	0.9156	0.8806	0.9639	0.9028
	FSIM	0.9442	0.9426	0.9136	0.9359	0.9561	0.9430	0.8986	0.9496	0.9373	0.9512	0.9372
NCSR ²⁰	PSNR	30.84	31.49	28.34	33.63	33.39	31.26	27.91	32.27	30.16	31.57	31.09
	SSIM	0.9476	0.8968	0.8591	0.8696	0.9354	0.9009	0.8304	0.9229	0.8704	0.9648	0.8998
	FSIM	0.9381	0.9371	0.9078	0.9333	0.9587	0.9389	0.9088	0.9551	0.9331	0.9508	0.9362
GSR ²¹	PSNR	29.88	31.69	27.78	34.45	32.83	31.47	28.26	31.61	30.19	30.59	30.88
	SSIM	0.9410	0.9046	0.8666	0.8826	0.9333	0.9135	0.8436	0.9165	0.8793	0.9575	0.9039
	FSIM	0.9218	0.9411	0.9006	0.9420	0.9574	0.9463	0.9155	0.9471	0.9349	0.9382	0.9345
Ours	PSNR	31.42	31.84	28.50	34.13	33.62	31.63	27.42	32.50	30.36	32.12	31.35
	SSIM	0.9525	0.9090	0.8770	0.8852	0.9373	0.9149	0.8244	0.9270	0.8811	0.9692	0.9078
	FSIM	0.9456	0.9442	0.9167	0.9368	0.9606	0.9465	0.9004	0.9564	0.9380	0.9580	0.9403

The bold values mean the largest ones. FISTA: fast iterative shrinkage/thresholding algorithm; PSNR: peak signal-to-noise; SSIM: structural similarity; FSIM: feature similarity; IDD-BM3D: iterative decoupled deblurring block-matching and 3D; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

Table 4. Comparison of the ISNR (dB) results of the deblurring methods in the second set.

Method	Scenario						Scenario					
	1	2	3	4	5	6	1	2	3	4	5	6
	Cameraman (256 × 256)						House (256 × 256)					
Input PSNR	22.23	22.16	20.76	24.62	23.36	29.82	25.61	25.46	24.11	28.06	27.81	29.98
IDD-BM3D ⁴⁰	8.85	7.12	10.45	3.98	4.31	4.89	9.95	8.55	12.89	5.79	5.74	7.13
NCSR ²⁰	8.78	6.69	10.33	3.78	4.60	4.50	9.96	8.48	13.12	5.81	5.67	6.94
GSR ²¹	8.39	6.39	10.08	3.33	3.94	4.76	10.02	8.56	13.44	6.00	5.95	7.18
Ours	9.01	7.68	11.39	4.07	4.64	5.16	10.07	8.31	13.58	5.85	5.98	7.24
	Lena (512 × 512)						Barbara (512 × 512)					
Input PSNR	27.25	27.04	25.84	28.81	29.16	30.03	23.34	23.25	22.49	24.22	23.77	29.78
IDD-BM3D ⁴⁰	7.97	6.61	8.91	4.97	4.85	6.34	7.64	3.96	6.05	1.88	1.16	5.45
NCSR ²⁰	8.03	6.54	9.25	4.93	4.86	6.19	7.76	3.64	5.92	2.06	1.43	5.50
GSR ²¹	8.24	6.76	9.43	5.17	4.96	6.57	8.98	4.80	7.15	2.19	1.58	6.20
Ours	8.32	6.89	9.56	5.30	5.02	6.69	9.21	3.60	5.83	1.93	1.55	6.10

The bold values mean the largest ones. ISNR: improvement of signal-to-noise ratio; PSNR: peak signal-to-noise; IDD-BM3D: iterative decoupled deblurring block-matching and 3D; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

Table 5. Comparison of the PSNR (dB) results of the deblurring methods with different similarity criteria.

9 × 9 Uniform blur with Gaussian noise $\sigma_n = \sqrt{2}$												
Images	Butterfly	Boats	C. Man	House	Parrot	Lena	Barbara	Starfish	Peppers	Leave	Avg.	
NCSR ²⁰	29.68	31.08	28.62	34.31	31.95	29.96	28.10	30.28	29.66	29.98	30.36	
Ours	30.60	31.40	28.71	34.61	32.57	30.21	27.96	30.85	29.94	30.86	30.77	
Ours-SSIM	30.48	31.40	28.64	34.57	32.40	30.08	27.90	30.62	29.81	30.59	30.65	
Ours-PCA	30.49	31.34	28.63	34.59	32.43	30.18	27.89	30.76	29.90	30.81	30.70	
Gaussian blur 1.6 with Gaussian noise $\sigma_n = \sqrt{2}$												
NCSR ²⁰	30.84	31.49	28.34	33.63	33.39	31.26	27.91	32.27	30.16	31.57	31.09	
Ours	31.42	31.84	28.50	34.13	33.62	31.63	27.42	32.50	30.36	32.12	31.35	
Ours-SSIM	31.23	31.74	28.42	34.09	33.43	31.50	27.34	32.36	30.24	31.89	31.22	
Ours-PCA	31.30	31.82	28.50	34.10	33.47	31.53	27.36	32.37	30.21	31.98	31.26	

The bold values mean the largest ones. PSNR: peak signal-to-noise; NCSR: nonlocally centralized sparse representation; SSIM: structural similarity; PCA: principal component analysis.

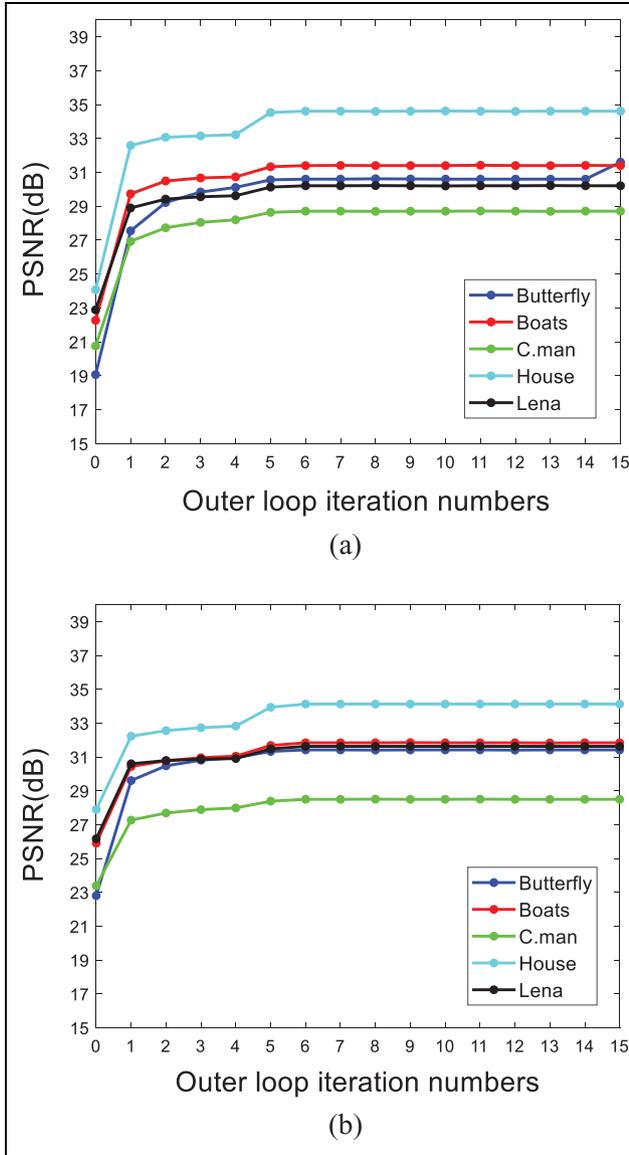


Figure 8. (a) The changing PSNR (dB) values of five test images as functions of iteration numbers for the uniform blur in the first set. (b) The changing PSNR (dB) values of five test images as functions of iteration numbers for the Gaussian blur in the first set. PSNR: peak signal-to-noise.

uniform blur kernels are shown in Figure 9. From Figure 9, it is concluded that the performance of our proposed model is not quite sensitive to K and Q because all the curves are almost flat. The highest performance for each image is usually achieved with K and Q in the range (60, 70) and (10, 20), respectively. Therefore, in this article, K and Q are empirically set to be 64 and 13, respectively.

Extensional experiments

In the previous section of this article, we conduct merely image deblurring via our proposed novel similarity criteria under the superior NCSR framework. To demonstrate the universality and extensibility of our novel similarity criteria,

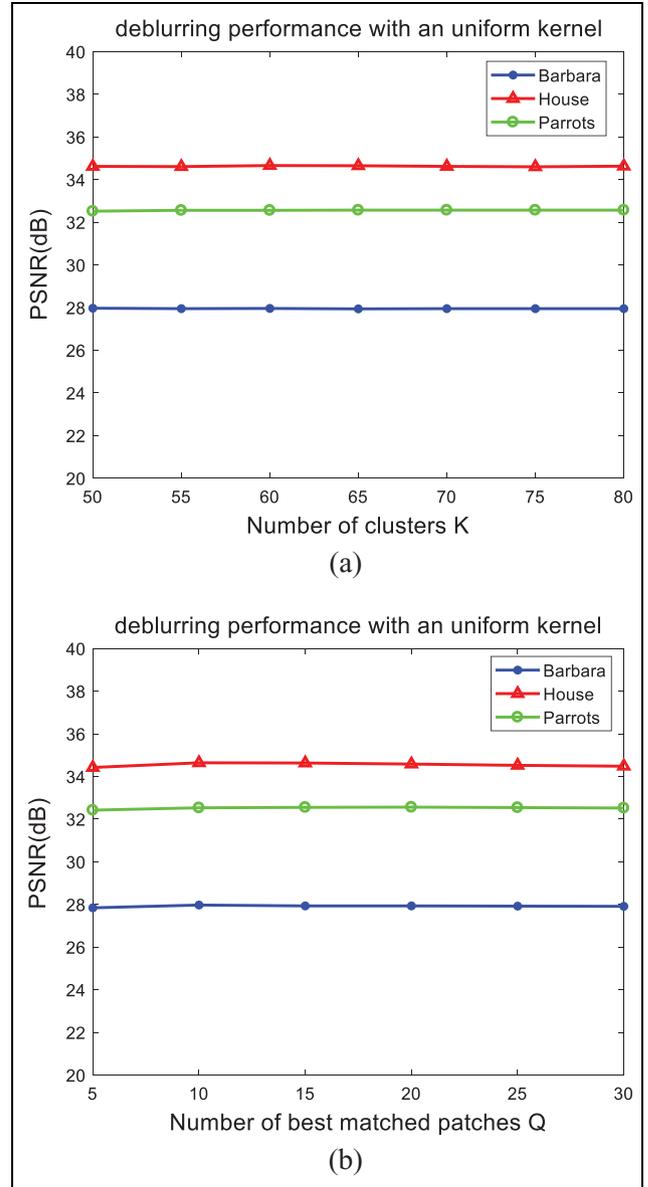


Figure 9. (a) Performance comparison with various K for three test images. (b) Performance comparison with various Q for three test images.

Table 6. The PSNR (dB) results of image super-resolution via our novel similarity criteria under the same NCSR framework.

9×9 Uniform blur with Gaussian noise $\sigma_n = \sqrt{2}$							
Images	Barbara	Boats	House	C. Man	Peppers	Lena	Avg.
GSR ²¹	28.95	31.34	34.48	28.28	29.66	30.10	30.47
Ours-GSR	29.10	31.30	34.65	28.49	29.78	30.34	30.61
Gaussian blur 1.6 with Gaussian noise $\sigma_n = \sqrt{2}$							
GSR ²¹	28.26	31.69	34.45	27.78	30.19	31.47	30.64
Ours-GSR	28.42	31.64	34.60	27.82	30.27	31.61	30.73

The bold values mean the largest ones. PSNR: peak signal-to-noise; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

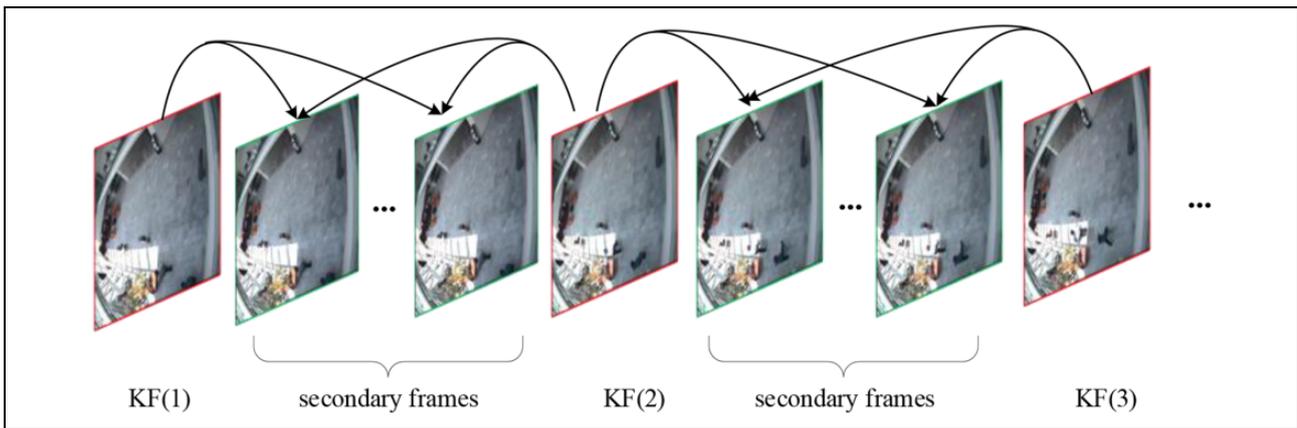
Table 7. The PSNR (dB) results of image deblurring via our novel similarity criteria under the GSR framework.

Images	Noiseless									
	Butterfly	Flower	Girl	Pathenon	Parrot	Raccon	Bike	Hat	Plants	Avg.
NCSR ²⁰	28.10	29.50	33.65	27.19	30.50	29.28	24.74	31.27	34.00	29.80
Ours-SR	28.49	29.65	33.68	27.32	30.44	29.25	24.88	31.38	34.16	29.92
Noisy, Gaussian noise of standard deviation is 5										
NCSR ²⁰	26.86	28.08	32.03	26.38	29.51	28.03	23.80	29.94	31.73	28.48
Ours-SR	27.39	28.23	32.03	26.60	29.51	28.00	24.00	30.15	31.92	28.65

The bold values mean the largest ones. PSNR: peak signal-to-noise; NCSR: nonlocally centralized sparse representation; GSR: group-based sparse representation.

Table 8. The concise procedure of video deblurring in real time.

- Step 1** Decompose the input video into L frames and divide the L frames into two parts, of which one is called key frames that consist of the frames $n \times t + 1$ and the other is called secondary frames that consist of the frames between key frames. We represent the $(n + 1)$ th key frames as $KF(n + 1)$, where $n = 0, 1, 2, \dots$ and $n \leq \frac{L-1}{t}$
- Step 2** Deblur each of the key frames $KF(n + 1)$ by our proposed algorithm (algorithm 1) and save certain information, that is, dictionaries of cluster $\mathbf{D} = \{\mathbf{D}_k\}_{k=1}^K$ and the index of location and weights of matched patches for each image patch
- Step 3** Simplify algorithm 1 in consideration of frame-to-frame correlation of a video: for each patch extracted from one secondary frame, we choose its subdictionary from subdictionaries of its nearest two key frames saved in step 2 by the way similar to the one described in algorithm 1 (Step (d)), and we can estimate its sparse coding coefficients using directly the index of location and weights of its front key frame
- Step 4** Deblur the secondary frames by simplified algorithm 1
- Step 5** Compose deblurred L frames into an output video

**Figure 10.** The process of video deblurring by our method.

we apply our novel similarity criteria in image super-resolution⁴³ under the same NCSR framework and image deblurring under another GSR framework,²¹ respectively.

The corresponding results are shown in Tables 6 and 7, respectively, which forcefully prove that our similarity criteria have advantageous influence on both NCSR super-resolution and GSR deblurring.

Application in real-time video deblurring

Though deblurring methods, for example, NCSR or GSR, which using nonlocal sparse regularization models, have achieved higher metrics objectively and favorable visual effects subjectively, they are all failed to be real time.

By analyzing the process of this kind of approach, we discover that the clustering and block matching processes introduced in algorithm 1 are so time-consuming that deblurring methods including the two processes can't deblur a single image in real time. However, considering frame-to-frame correlation of a video which consists of L frames in all, we can perform clustering and block matching processes for every T frames and try to apply our single image deblurring method to video deblurring. Because this is not our main content of research in this article, we introduce this idea simply in Table 8.

For a vivid illustration, we show the process in Figure 10. Note that if you expect a better performance of real-

time video deblurring, a further research is needed because of the complicated correlation among frames.

Conclusion and future work

This article presents a nonlocal sparse regularization model with novel similarity criteria called SSIM distance and PCA-subspace Euclidean distance for deblurring. The deblurring model we proposed is realized under the NCSR framework. The results of experiment on image deblurring have shown that the proposed method achieves significant performance improvements over many current state-of-the-art schemes. Extensional experiments forcefully prove the universality and extensibility of our similarity criteria in other deblurring framework and other image processing applications such as image super-resolution.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by National Natural Science Foundation of China (no 61501334).

ORCID iD

Wenxuan Shi  <http://orcid.org/0000-0002-2009-6404>

References

1. Sun L, Cho S, Wang J, et al. Edge-based blur kernel estimation using patch priors. In: *IEEE international conference on computational photography (ICCP)*, Cambridge, MA, USA, 19–21 April 2013, pp. 1–8. IEEE.
2. Dong J, Pan J, and Su Z. Blur kernel estimation via salient edges and low rank prior for blind image deblurring. *Signal Processing Image Commun* 2017; 58: 134–145.
3. Ribes A and Schmitt F. Linear inverse problems in imaging. *IEEE Signal Proc Mag* 2008; 25(4): 84–99.
4. Levin A, Weiss Y, Durand F, et al. Understanding and evaluating blind deconvolution algorithms. In: *IEEE international conference on computer vision and pattern recognition (CVPR)*, Miami, FL, USA, 20–25 June 2009, pp. 1964–1971. IEEE.
5. Cho S, Wang J, and Lee S. Handling outliers in non-blind image deconvolution. In: *IEEE international conference on computer vision (ICCV)*, Barcelona, Spain, 6–13 November 2011, pp. 495–502. IEEE.
6. Galatsanos NP and Katsaggelos AK. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Trans Image Proc* 1992; 1(3): 322–336.
7. Mumford D and Shah J. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun Pure Appl Math* 1989; 42: 577–685.
8. Bioucas-Dias JM. Bayesian wavelet-based image deconvolution: a GEM algorithm exploiting a class of heavy-tailed priors. *IEEE Trans Image Process* 2006; 15(4): 937–951.
9. Rudin LI, Osher S, and Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* 1992; 60(1–4): 259–268.
10. Deshpande AM and Patnaik S. Uniform and non-uniform single image deblurring based on sparse representation and adaptive dictionary learning. *Int J Multimed Appl* 2014; 6: 47.
11. Lee CC and Hwang WL. Sparse representation of a blur kernel for out-of-focus blind image restoration. In: *IEEE international conference on image processing (ICIP)*, Phoenix, Arizona, USA, 25–28 September 2016, pp. 2698–2702. IEEE.
12. Chen H, Wang Q, Wang C, et al. Image decomposition-based blind image deconvolution model by employing sparse representation. *IET Image Proc* 2016; 10(11): 908–925.
13. Starck JL, Nguyen MK, and Murtagh F. Wavelets and curvelets for image deconvolution: a combined approach. *Signal Proc* 2003; 83: 2279–2283.
14. Rubinstein R, Bruckstein AM, and Elad M. Dictionaries for sparse representation modeling. *Proc IEEE* 2010; 98(6): 1045–1057.
15. Daubechies I, Defrise M, and De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun Pure Appl Math* 2004; 57(11): 1413–1457.
16. Peyré G. Image processing with nonlocal spectral bases. *Mult Model Simulat* 2008; 7(2): 703–730.
17. Zhang X, Burger M, Bresson X, et al. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM J Imaging Sci* 2010; 3(3): 253–276.
18. Jung M, Bresson X, Chan TF, et al. Nonlocal Mumford-Shah regularizers for color image restoration. *IEEE Trans Image Process* 2011; 20(6): 1583–1598.
19. Dong W, Shi G, Li X, et al. Image reconstruction with locally adaptive sparsity and nonlocal robust regularization. *Signal Processing: Image Commun* 2012; 27: 1109–1122.
20. Dong W, Zhang L, Shi G, et al. Nonlocally centralized sparse representation for image restoration. *IEEE Trans Image Proc* 2013; 22(4): 1620–1630.
21. Zhang J, Zhao D, and Gao W. Group-based sparse representation for image restoration. *IEEE Trans Image Proc* 2014; 23(8): 3336–3351.
22. Liu S, Wu G, Liu H, et al. Image restoration approach using a joint sparse representation in 3D-transform domain. *Digital Signal Proc* 2017; 60: 307–323.
23. Aggarwal CC and Reddy CK. *Data clustering: algorithms and applications*. 2013; 29(10): 149A. England: Chapman & Hall/CRC.
24. Qi J, Yu Y, Wang L, et al. An effective and efficient hierarchical K-means clustering algorithm. *Int J Distrib Sens Net* 2017; 13(8): 1–17.

25. Zhang Y, Liu J, Li M, et al. Joint image denoising using adaptive principal component analysis and self-similarity. *Inf Sci* 2014; 259: 128–141.
26. Tombari F and Di Stefano L. Bounded non-local means for fast and effective image denoising. In: *International conference on image analysis and processing (ICIAP)*, Genova, Italy, 7–11 September 2015, pp. 183–193.
27. Wang Z, Bovik AC, Sheikh HR, et al. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Proc* 2004; 13(4): 600–612.
28. Zhang L, Zhang L, Mou X, et al. FSIM: a feature similarity index for image quality assessment. *IEEE Trans Image Process* 2011; 20(8): 2378–2386.
29. Schuler CJ, Christopher Burger H, Harmeling S, et al. A machine learning approach for non-blind image deconvolution. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Portland, OR, USA, 23–28 June 2013, pp. 1067–1074. IEEE.
30. Xu L, Ren JS, Liu C, et al. Deep convolutional neural network for image deconvolution. In: *International conference on neural information processing systems (NIPS)*, Montréal, Canada, 8–13 December 2014, pp. 1790–1798.
31. Sun J, Cao W, Xu Z, et al. Learning a convolutional neural network for non-uniform motion blur removal. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, Boston, 8–10 June 2015, pp. 769–777. IEEE.
32. Jin KH, McCann MT, Froustey E, et al. Deep convolutional neural network for inverse problems in imaging. *IEEE Trans Image Proc* 2017; 26: 4509–4522.
33. Huang J and Mumford D. Statistics of natural images and models. In: *IEEE computer society conference on computer vision and pattern recognition (CVPR)*, Fort Collins, CO, USA, 23–25 June 1999, pp. 541–547. IEEE.
34. Lee AB, Pedersen KS, and Mumford D. The nonlinear statistics of high-contrast patches in natural images. *Int J Comput Vis* 2003; 54(1–3): 83–103.
35. Muresan DD and Parks TW. Adaptive principal components and image denoising. In: *Proceedings of the IEEE international conference on image processing (ICIP)*, Barcelona, Spain, 14–17 September 2003, pp. 1–101-4. IEEE.
36. Meng J and Yang Y. Symmetrical two-dimensional PCA with image measures in face recognition. *Int J Adv Robot Syst* 2012; 238(9): 1–10.
37. James R, Jolly AM, Anjali C, et al. Image denoising using adaptive PCA and SVD. In: *FIFTH international conference on advances in computing and communications (ICACC)*, Kochi, Kerala, India, 3–5 September 2015, pp. 383–386. IEEE.
38. Klema V and Laub A. The singular value decomposition: its computation and some applications. *IEEE Trans Automat Control* 1980; 25(2): 164–176.
39. Portilla J. Image restoration through l0 analysis-based sparse optimization in tight frames. In: *16th IEEE international conference on image processing (ICIP)*, Cairo, Egypt, 7–12 November 2009, pp. 3909–3912. IEEE.
40. Danielyan A, Katkovnik V, and Egiazarian K. BM3D frames and variational image deblurring. *IEEE Trans Image Proc* 2012; 21(4): 1715–1728.
41. Beck A and Teboulle M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans Image Proc* 2009; 18(11): 2419–2434.
42. Dabov K, Foi A, Katkovnik V, et al. Image restoration by sparse 3D transform-domain collaborative filtering. In: *Image processing: algorithms and systems VI*, San Jose, California, USA, January 2008, p. 681207. SPIE.
43. Zhu Z, Liu X, Yuan Y, et al. Superresolution reconstruction of video sequence using a coarse-to-fine registration and optimal interpolation strategy. *Int J Adv Robot Systems* 2013; 245(10): 1–10.