

# Centralized Sparse Representation for Image Restoration

Weisheng Dong

Sch. of Elec. Engineering  
Xidian University, China

wsdong@mail.xidian.edu.cn

Lei Zhang

Dept. of Computing  
The Hong Kong Polytechnic Univ.

cslzhang@comp.polyu.edu.hk

Guangming Shi

Sch. of Elec. Engineering  
Xidian University, China

gmshi@xidian.edu.cn

## Abstract

*This paper proposes a novel sparse representation model called centralized sparse representation (CSR) for image restoration tasks. In order for faithful image reconstruction, it is expected that the sparse coding coefficients of the degraded image should be as close as possible to those of the unknown original image with the given dictionary. However, since the available data are the degraded (noisy, blurred and/or down-sampled) versions of the original image, the sparse coding coefficients are often not accurate enough if only the local sparsity of the image is considered, as in many existing sparse representation models. To make the sparse coding more accurate, a centralized sparsity constraint is introduced by exploiting the nonlocal image statistics. The local sparsity and the nonlocal sparsity constraints are unified into a variational framework for optimization. Extensive experiments on image restoration validated that our CSR model achieves convincing improvement over previous state-of-the-art methods.*

## 1. Introduction

Image restoration (IR) aims to recover a high-quality image from its degraded (e.g., noisy, blurred and/or down-sampled) versions, which may be taken, for example, by a low-end camera and/or under limited conditions. For an observed image  $\mathbf{y}$ , the problem of IR can be expressed by

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

where  $\mathbf{H}$  is a degradation matrix,  $\mathbf{x}$  is the original image vector and  $\mathbf{v}$  is the additive noise vector. Due to the ill-posed nature of IR, the regularization techniques, which try to incorporate both the observation model and the prior knowledge of the desired solution into a variational formulation, have been extensively studied [4]. For regularization methods, finding and modeling the appropriate prior knowledge of natural images is one of the most important concerns, and hence various methods have been proposed to learn the prior knowledge of natural images [25, 5, 6, 12].

In recent years the sparse representation based modeling has been proven to be a promising model for image restoration [9, 5, 13, 20, 16, 21, 27, 15, 14]. In the study of human visual system [23, 24], it has been found that cell receptive fields code natural images using a small number of structural primitives sparsely chosen out from an over-complete code set. Mathematically, a signal  $\mathbf{x} \in R^N$  can be represented as a linear combination of a few atoms from a dictionary  $\Phi$ , i.e.,  $\mathbf{x} \approx \Phi\alpha$ , via  $l_0$ -minimization:

$$\alpha_{\mathbf{x}} = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_0, \text{ s.t. } \|\mathbf{x} - \Phi\alpha\|_2 < \epsilon, \quad (2)$$

where  $\|\cdot\|_0$  counts the number of the non-zero coefficients in  $\alpha$  and  $\epsilon$  is a small constant balancing the sparsity and the approximation error. In practice, the  $l_0$ -norm is often replaced by the  $l_1$ -norm for efficient convex optimization [13, 9]. On the other hand, compared with the analytically designed dictionary (e.g., wavelet dictionary), the learned dictionary [1, 21, 19] from example image patches can better adapt to the signal  $\mathbf{x}$  and characterize the image structures.

In the scenario of IR, we have some degraded observation  $\mathbf{y}$  of the original image  $\mathbf{x}$ , i.e.,  $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$ . To reconstruct  $\mathbf{x}$  from  $\mathbf{y}$ , first  $\mathbf{y}$  is sparsely coded over  $\Phi$  by solving the following minimization problem:

$$\alpha_{\mathbf{y}} = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_1, \text{ s.t. } \|\mathbf{y} - \mathbf{H}\Phi\alpha\|_2 < \epsilon, \quad (3)$$

and then the reconstructed  $\mathbf{x}$ , denoted by  $\hat{\mathbf{x}}$ , is obtained as  $\hat{\mathbf{x}} = \Phi\alpha_{\mathbf{y}}$ . Clearly,  $\alpha_{\mathbf{y}}$  is expected to be very close to  $\alpha_{\mathbf{x}}$  computed in Eq. (3) so that the estimated image  $\hat{\mathbf{x}}$  can be very close to the true image  $\mathbf{x}$ . Unfortunately, since  $\mathbf{y}$  is noise corrupted, blurred and/or incomplete, the coding vector  $\alpha_{\mathbf{y}}$  resolved by Eq. (3) may deviate much from the desired vector  $\alpha_{\mathbf{x}}$ , leading to an inaccurate restoration of the image  $\mathbf{x}$ . In other words, the model in Eq. (3) can ensure  $\alpha_{\mathbf{y}}$  being sparse but cannot ensure  $\alpha_{\mathbf{y}}$  being as close to  $\alpha_{\mathbf{x}}$  as possible.

In this paper, we introduce the concept of sparse coding noise (SCN) to facilitate the discussion of the problem. The

SCN of  $\mathbf{y}$  is defined as

$$\mathbf{v}_\alpha = \alpha_x - \alpha_y. \quad (4)$$

We can see that given the dictionary  $\Phi$ , the IR result depends on the level of SCN  $\mathbf{v}_\alpha$  because the image reconstruction error  $\mathbf{v}_x = \hat{\mathbf{x}} - \mathbf{x} \approx \Phi\alpha_y - \Phi\alpha_x = \Phi\mathbf{v}_\alpha$ . The definition of SCN  $\mathbf{v}_\alpha$  also indicates one way to improve the quality of IR, that is, reduce the level of  $\mathbf{v}_\alpha$ .

The conventional sparse representation model such as that in Eq.(2) or Eq.(3) exploits mainly the (patch based) local sparsity of natural images. Each patch is coded individually without considering other patches. Nonetheless, the sparse coding coefficients  $\alpha$  are not randomly distributed because the image local patches are non-locally correlated. The nonlocal means (NLM) methods, which aim to exploit the image nonlocal redundancy, have been successfully used in many image processing applications, particularly in denoising [6]. This implies that the SCN can be reduced by exploiting the image nonlocal similarity, and hence the restored image quality can be improved. Indeed, some recent works, such as [10] and [20], are based on such considerations. For example, in [20] a group sparse coding scheme was proposed to code similar patches simultaneously, and it achieves impressive denoising results.

In this paper, we propose a centralized sparse representation (CSR) model to effectively reduce the SCN and thus enhance the sparsity based IR performance. The basic idea is to integrate the image *local sparsity constraint* (i.e., a local patch can be coded by a few atoms sparsely selected from a dictionary) and the *centralized sparsity constraint* (i.e., the sparse coding coefficients should be close to their mean values) into a unified variational framework for optimization. Specifically, in addition to requiring that the coding coefficients of each local patch are sparse, we also enforce the sparse coding coefficients to have small SCN via exploiting the nonlocal similarity induced sparsity, which can be characterized by the  $l_1$ -norm. Extensive experiments on IR are conducted, and the experimental results show that the proposed CSR algorithm outperforms significantly many state-of-the-art IR methods.

## 2. Centralized sparse representation modeling

### 2.1. The sparse coding noise in image restoration

Following the notation used in [16], we denote by  $\mathbf{x} \in R^N$  the original image, and by  $\mathbf{x}_i = \mathbf{R}_i\mathbf{x}$  an image patch of size  $\sqrt{n} \times \sqrt{n}$  at location  $i$ , where  $\mathbf{R}_i$  is the matrix extracting patch  $\mathbf{x}_i$  from  $\mathbf{x}$  at location  $i$ . Given a dictionary  $\Phi \in R^{n \times M}$ ,  $n < M$ , each patch can be sparsely coded as  $\mathbf{x}_i \approx \Phi\alpha_i$  by using some sparse coding algorithm [13, 9]. Then the entire image  $\mathbf{x}$  can be sparsely represented by the set of sparse codes  $\{\alpha_i\}$ . If the patches are allowed to

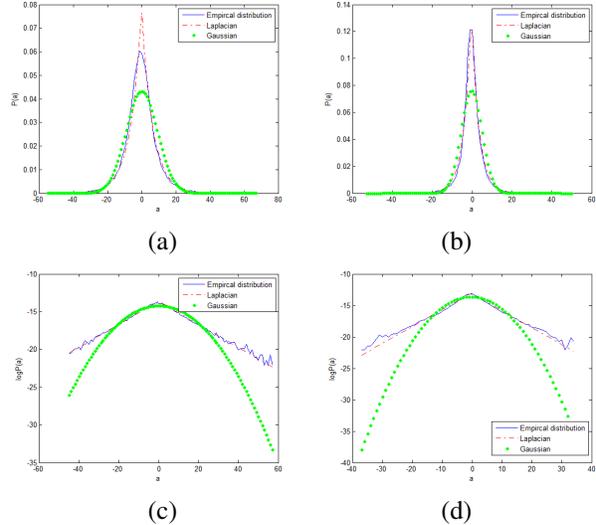


Figure 1. The distribution of SCN when the Lena image (a) is noisy and blurred; and (b) is down-sampled. (c) and (d) show the same distributions of (a) and (b) in log domain, respectively.

be overlapped, we obtain a very redundant patch-based image representation. Reconstructing  $\mathbf{x}$  from the sparse codes  $\{\alpha_i\}$  is then an over-determined system, and a straightforward least-square solution is [16]:

$$\mathbf{x} \approx \Phi \circ \alpha_x \triangleq \left( \sum_{i=1}^N \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \sum_{i=1}^N (\mathbf{R}_i^T \Phi \alpha_i), \quad (5)$$

where  $\alpha_x$  denotes the concatenation of all  $\alpha_i$ . The above equation is nothing but telling that the overall image is reconstructed by averaging each reconstructed patch of  $\mathbf{x}_i$ .

In the application of IR,  $\mathbf{x}$  is not available to code and what we have is the degraded observation  $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$ . The sparse coding of  $\mathbf{x}$  is based on  $\mathbf{y}$  via minimizing

$$\alpha_y = \operatorname{argmin}_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 \}. \quad (6)$$

The image is then reconstructed as  $\hat{\mathbf{x}} = \Phi \circ \alpha_y$ . As we defined and discussed in Eq. (4), coefficients  $\alpha_y$  will deviate from  $\alpha_x$ , and the sparse coding noise (SCN)  $\mathbf{v}_\alpha = \alpha_y - \alpha_x$  determines the IR quality of  $\hat{\mathbf{x}}$ .

Here we perform some experiments to investigate the statistics of the SCN  $\mathbf{v}_\alpha$ . We use the image *Lena* as an example. The original image  $\mathbf{x}$  is first blurred (by a Gaussian blur kernel with standard deviation 1.6), and Gaussian white noise of standard deviation  $\sqrt{2}$  is added to get a noisy and blurred image  $\mathbf{y}$ . Then we compute  $\alpha_x$  and  $\alpha_y$  by minimizing

$$\alpha_x = \operatorname{argmin}_{\alpha} \{ \|\mathbf{x} - \Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 \}. \quad (7)$$

and Eq. (6), respectively. The DCT dictionary is adopted in this experiment. Then the SCN is computed as  $\mathbf{v}_\alpha =$

$\alpha_x - \alpha_y$ . In Fig. 1(a), we plot the distribution of  $v_\alpha$  corresponding to the 4<sup>th</sup> atom (other atoms exhibit similar distributions) in the dictionary. In Fig. 1(b), we plot the distribution of  $v_\alpha$  when the observed data  $\mathbf{y}$  is first blurred (by a Gaussian blur kernel with standard deviation 1.6) and then down-sampled. We can see that the empirical distributions of the SCN  $v_\alpha$  are highly peaked at zero and can be well characterized by Laplacian functions, while Gaussian functions have much bigger fitting error. In Figs. 1(c) and (d) we show these distributions in log domain to better observe the fitting of the tails. This observation motivates us to model the SCN with a Laplacian prior, as will be introduced in Section 3.

## 2.2. Centralized sparse representation

It is apparent that suppressing the SCN  $v_\alpha$  could improve the IR output  $\hat{x}$ . However, the difficulty lies in that the coefficient vector  $\alpha_x$  is unknown so that  $v_\alpha$  cannot be directly measured. Nonetheless, if we could have some reasonably good estimation of  $\alpha_x$ , denoted by  $\hat{\alpha}_x$ , available, then  $\alpha_y - \hat{\alpha}_x$  can be an estimation of SCN  $\alpha_x$ . Intuitively, to suppress  $\alpha_x$  and improve the accuracy of  $\alpha_y$ , a new sparse coding model can be

$$\alpha_y = \operatorname{argmin}_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 + \gamma \|\alpha - \hat{\alpha}_x\|_{l_p} \}. \quad (8)$$

where  $\gamma$  is a constant and the  $l_p$ -norm ( $p$  can be 1 or 2) is used to measure the distance between  $\alpha$  and  $\hat{\alpha}_x$ . Compared with Eq. (6), Eq. (8) enforces  $\alpha_y$  being close to  $\hat{\alpha}_x$  (so that the SCN  $v_\alpha$  can be suppressed) while enforcing  $\alpha_y$  being sparse, and hence the resulted sparse code will be more desirable than that by solving Eq. (6). If  $\hat{\alpha}_x = 0$  and  $p = 1$ , model in Eq. (8) will be reduced to the conventional model in Eq. (6).

Now the problem turns to how to find a reasonable estimation of the unknown vector  $\alpha_x$ . By viewing  $\alpha_x$  as a random variable vector, one good unbiased estimation of  $\alpha_x$  is naturally the mean of it; that is, we could set  $\hat{\alpha}_x = E[\alpha_x]$ . In practice, we could approximate  $E[\alpha_x]$  by  $E[\alpha_y]$  by assuming that the SCN  $v_\alpha$  is nearly zero mean (please refer to the empirical observations in Section 2.1), and there is  $\hat{\alpha}_x = E[\alpha_x] \approx E[\alpha_y]$ . Then Eq. (8) can be converted into

$$\alpha_y = \operatorname{argmin}_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 + \gamma \|\alpha - E[\alpha]\|_{l_p} \}. \quad (9)$$

We call the above model *centralized sparse representation* (CSR) because it enforces the sparse coding coefficient  $\alpha$  to approach to its distribution center (i.e. the mean value).

For the sparse code  $\alpha_i$  on each image patch  $i$ ,  $E[\alpha_i]$  can be nearly computed if we could have enough samples of  $\alpha_i$ . Fortunately, in natural images there are often many nonlocal similar patches to the given patch  $i$ . Then  $E[\alpha_i]$  can be computed as the weighted average of those sparse

code vectors associated with the nonlocal similar patches (including patch  $i$ ) to patch  $i$ . To this end, we can form a cluster, denoted by  $C_i$ , for each patch  $i$  via block matching and then average the sparse codes within each cluster.

Denote by  $\alpha_{i,j}$  the sparse code of the searched similar patch  $j$  to patch  $i$ . Then we use the weighted average of all  $\alpha_{i,j}$  to approximate  $E[\alpha_i]$ . There is

$$\mu_i = \sum_{j \in C_i} \omega_{i,j} \alpha_{i,j}, \quad (10)$$

where  $\omega_{i,j}$  is the weight. Like that in nonlocal means methods [6],  $\omega_{i,j}$  can be set to be inverse proportional to the distance between patches  $i$  and  $j$ :

$$\omega_{i,j} = \exp(-\|\hat{x}_i - \hat{x}_{i,j}\|_2^2/h)/W, \quad (11)$$

where  $\hat{x}_i = \Phi \hat{\alpha}_i$  and  $\hat{x}_{i,j} = \Phi \hat{\alpha}_{i,j}$  are the estimates of patches  $i$  and  $j$ ,  $W$  is a normalization factor and  $h$  is a pre-determined scalar.

By taking  $\mu_i$  as the approximation of  $E[\alpha_i]$ , the CSR model in Eq. (9) can be written as:

$$\alpha_y = \operatorname{argmin}_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 + \gamma \sum_{i=1}^N \|\alpha_i - \mu_i\|_{l_p} \}. \quad (12)$$

From Eq. (12) we can more clearly see that the CSR model unifies the local sparsity (i.e.  $\|\alpha\|_1$ ) and nonlocal similarity induced sparsity (i.e.  $\|\alpha_i - \mu_i\|_{l_p}$ ) into a unified variational formulation. By exploiting both the local and nonlocal redundancy, better IR results can be expected.

Eq.(12) actually indicates an iterative minimization approach to the CSR model. We initialize  $\mu_i$  as  $\mathbf{0}$ , i.e.,  $\mu_i^{(-1)} = \mathbf{0}$ . Then from some initial sparse coding result, denoted by  $\alpha_y^{(0)}$ , we can get the initial estimation of  $x$ , denoted by  $x^{(0)}$ , via  $x^{(0)} = \Phi \circ \alpha_y^{(0)}$ . Based on  $x^{(0)}$  we can find the similar patches of each local patch  $i$ , and hence the non-local mean of the coding vector of each patch, i.e.  $\mu_i$ , can be updated by  $\alpha_y^{(0)}$  via Eqs. (10) and (11). The updated mean value, denoted by  $\mu_i^{(0)}$ , will be used in the next round of sparse coding process with the CSR model. Such a procedure is iterated until convergence. In the  $j^{th}$  iteration, the sparse coding is performed by

$$\alpha_y^{(j)} = \operatorname{argmin}_{\alpha} \{ \|\mathbf{y} - \mathbf{H}\Phi \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 + \gamma \sum_{i=1}^N \|\alpha_i - \mu_i^{(j-1)}\|_{l_p} \}. \quad (13)$$

From the above discussion, it can be seen that the sparse coding and non-local clustering steps are alternatively implemented in the proposed CSR scheme. During the iteration, the accuracy of sparse code  $\alpha_y^{(j)}$  is gradually improved, which improves the accuracy of non-local clustering in return, and the improvement of non-local clustering

further improves the accuracy of sparse coding. Finally, the desired sparse code  $\alpha_y$  is obtained when the joint sparse coding and non-local clustering process falls into a local minimum. Please note that the model in Eq. (12) is not convex but it is convex when the mean  $\mu_i$  is fixed. That is, the sparse coding step in Eq. (13) is convex when  $p \geq 1$ .

### 3. Algorithm of CSR

#### 3.1. The determination of parameters $\lambda$ and $\gamma$

In the proposed CSR model, there are two parameters,  $\lambda$  and  $\gamma$ , which balance the local redundancy induced sparsity and the nonlocal redundancy induced sparsity, respectively. These two parameters can be set empirically using a training set. However, a more reasonable and adaptive setting of them could not only improve the convergence speed but also improve much the IR quality [7]. In this sub-section, we provide a Bayesian interpretation of the sparse coding step of the CSR model, and this also presents us an explicit way to determine the parameters  $\lambda$  and  $\gamma$ . Actually, in wavelet based image denoising [26], the connection between wavelet representation and Bayesian framework has been well established. Such a connection helps to reconcile the difference between the deterministic and probabilistic approaches.

From the discussions in Section 2.2, we know that the CSR model will yield two outputs, the sparse code  $\alpha_y$  and the mean  $\mu_i$ , whereas we only have interest in the former one because the image is restored from  $\alpha_y$ . In other words,  $\mu_i$  or  $E[\alpha_y]$  is a hidden variable used in the optimization process. For the convenience of following development, we let  $\theta = \alpha - E[\alpha]$ . Then given the observation  $y$ , the maximum a posteriori (MAP) estimation of  $\alpha$  and  $\theta$  can be formulated as:

$$(\alpha_y, \theta_y) = \underset{\alpha, \theta}{\operatorname{argmax}} P(\alpha, \theta | y). \quad (14)$$

According to the Bayesian formula, we can have

$$(\alpha_y, \theta_y) = \underset{\alpha, \theta}{\operatorname{argmax}} \{P(y | \alpha, \theta) \times P(\alpha, \theta)\}, \quad (15)$$

where the two terms correspond to the likelihood and prior terms, respectively. Following the observation model of Eq. (1), the likelihood term can be characterized as:

$$P(y | \alpha, \theta) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{1}{2\sigma_n^2} \|y - \mathbf{H}\Phi \circ \alpha\|_2^2\right), \quad (16)$$

where  $\sigma_n$  is the standard deviation of the additive Gaussian noise.

It can be empirically found that  $\alpha$  and  $\theta$  are nearly uncorrelated. For instance, for the nine images used in our deblurring experiments (refer to Section 4.1 please), the correlation coefficients between  $\alpha$  and  $\theta$  range from 0.039 to

0.153. Therefore, in this paper we assume that  $\alpha$  and  $\theta$  are independent of each other and both  $\alpha$  and  $\theta$  are i.i.d. random variable vectors. As shown in Fig. 1, the SCN can be well characterized by the Laplacian distribution. Thus, we can model the SCN signal  $\theta$  with i.i.d Laplacian distributions. Meanwhile, it is well accepted in literature that the sparse coefficients  $\alpha$  can be characterized by i.i.d Laplacian distributions. Hence, the prior term in Eq. (15) can be expressed as

$$P(\alpha, \theta) = \prod_i \frac{1}{\sqrt{2}\sigma_i} \exp\left(-\frac{|\alpha_i|}{\sigma_i}\right) \times \prod_i \frac{1}{\sqrt{2}\delta_i} \exp\left(-\frac{|\theta_i|}{\delta_i}\right), \quad (17)$$

where  $\alpha_i$  and  $\theta_i$  are the  $i^{\text{th}}$  elements of  $\alpha$  and  $\theta$ , respectively, and  $\sigma_i$  and  $\delta_i$  are the standard deviations of  $\alpha_i$  and  $\theta_i$ , respectively.

Substituting Eqs. (16) and (17) into Eq. (15), we can obtain

$$\begin{aligned} \alpha_y = \underset{\alpha}{\operatorname{argmin}} \{ & \|y - \mathbf{H}\Phi \circ \alpha\|_2^2 + \sum_i \frac{2\sqrt{2}\sigma_n^2}{\sigma_i} \|\alpha_i\|_1 \\ & + \sum_i \frac{2\sqrt{2}\sigma_n^2}{\delta_i} \|\theta_i\|_1 \}. \end{aligned} \quad (18)$$

By comparing Eq. (18) with Eq. (9), it is obvious that the  $l_1$ -norm (i.e.  $p = 1$ ) should be chosen to characterize the centralized sparsity term so that the Bayesian optimal estimation of  $\alpha$  can be achieved. This is simply because the distribution of  $\theta$  can be well modeled as Laplacian. Therefore, the CSR model in Eq. (9) can be specified as

$$\alpha_y = \underset{\alpha}{\operatorname{argmin}} \{ \|y - \mathbf{H}\Phi \circ \alpha\|_2^2 + \sum_i \lambda_i \|\alpha_i\|_1 + \sum_i \gamma_i \|\theta_i\|_1 \}. \quad (19)$$

Finally, comparing Eq. (18) and Eq. (19), we have

$$\lambda_i = \frac{2\sqrt{2}\sigma_n^2}{\sigma_i}, \quad \gamma_i = \frac{2\sqrt{2}\sigma_n^2}{\delta_i}. \quad (20)$$

In implementation,  $\sigma_i$  and  $\delta_i$  are estimated from the sets of  $\alpha_i$  and  $\theta_i$  from the collected nonlocal similar patches. This estimation is more robust than those using only local patches.  $\lambda_i$  and  $\gamma_i$  are then updated with the update of  $\alpha$  and  $\theta$  in each iteration (or in several iterations).

#### 3.2. The selection of dictionary $\Phi$

In previous sections, we present the CSR model and the associated algorithms by supposing that dictionary  $\Phi$  is given. The proposed CSR is a general model and the selection of dictionary  $\Phi$  can be various. For example, the wavelet dictionary can be used, or a learned dictionary from example images by using algorithms such as KSVD [1] can

be used. However, the analytically designed wavelet dictionary and the learned KSVD dictionary are universal dictionaries; that is, they can be used to represent any image patch but they may lack sufficient flexibility to sparsely represent a given local patch. In the proposed CSR algorithms, the many similar patches to a given one are collected. This motivates us to use an adaptive dictionary to each patch. Instead of learning a universal dictionary, we learn a local dictionary for each patch or each cluster of similar patches. Specifically, we apply PCA to each cluster to learn a dictionary of PCA bases and use this dictionary to code the patches in this cluster. Such a local PCA strategy has been used in [28, 8, 15, 14] for image reconstruction.

### 3.3. Summary of the algorithm

From the analysis in previous sections, we can see that the CSR model in Eq. (9) or Eq. (12) can be iteratively solved. From some initialization, once the sparse code  $\alpha_y$  is solved, the nonlocal means  $\mu_i$  can be calculated, and then the code  $\alpha_y$  can be updated by Eq. (19), and so on. Eq. (19) has two  $l_1$ -norm constraints, and it can be generally solved by using the Augmented Lagrange Multiplier (ALM) methods [3]. In this paper, we extended the iterative shrinkage (IS) algorithm in [13] from one  $l_1$ -constraint to two  $l_1$ -constraints. Although the IS algorithm needs many iterations for convergence, it is very simple within each iteration. More importantly, the parameters  $\lambda_i$  and  $\gamma_i$  in Eq. (20) can be directly converted into the thresholds in the IS algorithm. Due to the limit of space, we omit the details here and interested readers may refer to [13] for more details.

The main procedures of the proposed CSR based image restoration algorithm are summarized in Algorithm 1.

#### Algorithm 1. Image Restoration with CSR

- Initialization: Compute an initial estimate  $\hat{x}$  using the standard sparse model [13];
- Outer loop: iterate on  $l = 1, 2, \dots, L$ 
  - Update the dictionary for each cluster of similar patches by using PCA;
  - Update the regularization parameters ( $\lambda$  and  $\gamma$ ) using Eq. (20);
  - Calculate the nonlocal means  $\mu_i^{(l-1)}$  from the sparse codes  $\alpha_y^{(l-1)}$ ;
  - Calculate  $\alpha_y^{(l)}$  by solving Eq. (19) via the extended iterative shrinkage algorithm [13].

## 4. Experimental results

We conduct extensive experiments on IR to demonstrate the performance of the proposed CSR model. On image denoising, CSR could achieve very similar results to BM3D [10] and the group sparse coding method [20]. Due to the

limit of space, we only show the results on image deblurring and super-resolution in this section. The patch size in our implementation is  $6 \times 6$ . The source code and more experimental results of the proposed CSR approach can be found at <http://www.comp.polyu.edu.hk/~cslzhang/CSR.htm>.

### 4.1. Image deblurring

The deblurring performance of the CSR model was verified both on simulated blur images and real blur images. To simulate a blur image, the original images were blurred by a blur kernel and then additive Gaussian noise with standard deviations  $\sigma_n = \sqrt{2}$  and  $\sigma_n = 2$  were added. Two blur kernels, a  $9 \times 9$  uniform kernel and a Gaussian blur kernel with standard deviation 1.6, were used for simulation. For the real motion blurred images, we borrowed the kernel estimation method from [17] to estimate the blur kernels.

We compared the CSR deblurring approach to several recently developed deblurring methods, i.e., the constrained TV deblurring (denoted by FISTA) method [2], the SA-DCT deblurring method [18], and the BM3D deblurring method [11]. Note that FISTA is a recently proposed TV-based deblurring approach that can well reconstruct the piecewise smooth regions. The SA-DCT and BM3D are two famous image restoration methods that often produce state-of-the-art image deblurring results.

The PSNR results on a set of 9 photographic images are reported in Table 1. From Table 1, we can conclude that the proposed CSR deblurring method significantly outperforms other competing methods for both uniform blurring and Gaussian blurring. The visual comparisons of the deblurring methods are shown in Fig. 2, from which we can see that the CSR model leads to much cleaner and sharper image edges and textures than other methods. More examples can be found in the supplementary materials.

We also applied the proposed CSR model to some real motion blurred images, where the real blurring kernels are unknown. Since the blurring kernel estimation is a non-trivial task and it is out of the scope of this paper, we borrowed the blur kernel estimation method from [17] to estimate the blur kernels. The estimated blur kernels were then fed into the proposed CSR approach. In Fig. 3, we presented the deblurred images of a real blurring image by the blind deblurring method of [17] and the proposed CSR. We can see that the image restored by our approach is much clearer and much more details are recovered.

### 4.2. Single image super-resolution

In single image super-resolution, the observed low-resolution (LR) image is obtained by first blurring with a blur kernel and then downsampling by a scaling factor. Hence, recovering the high-resolution (HR) image from a single LR image is more underdetermined than image deblurring. In this subsection, we conducted experiments

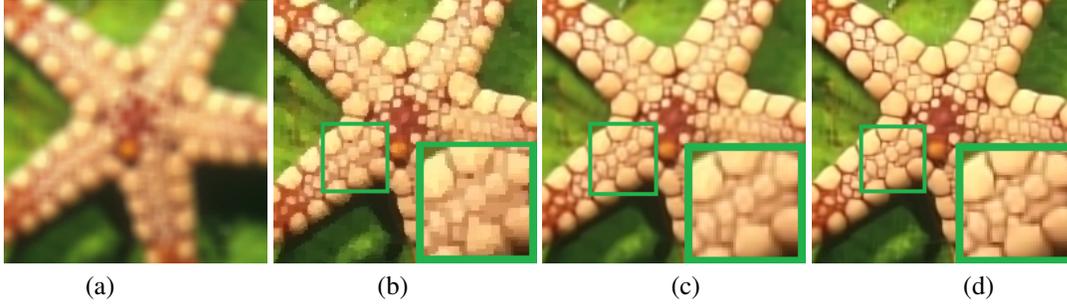


Figure 2. Image deblurring performance comparison for *Starfish* image ( $9 \times 9$  uniform blur,  $\sigma_n = \sqrt{2}$ ). (a) Noisy and blurred; (b) FISTA [2] (PSNR=27.75 dB); (c) BM3D [11] (PSNR=28.61dB); (d) the proposed CSR (PSNR=30.30 dB).

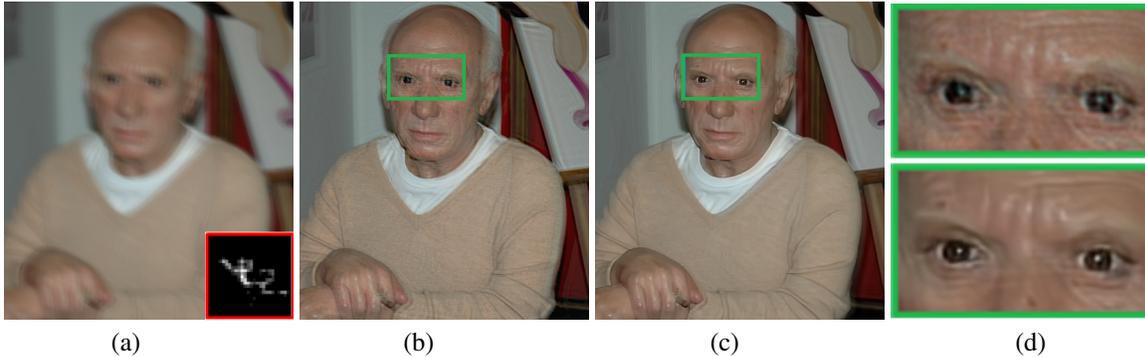


Figure 3. Deblurring performance comparison for a real blurred image with the blurring kernel estimated using the kernel estimation approach from [17]. (a) Input blurred image; (b) deblurred image by [17]; (c) the deblurred image by CSR; (d) close-up views.

of single image super-resolution using the proposed CSR method and other competing methods. The observed LR image is generated by first blurring an HR image with a blur kernel, i.e., a  $7 \times 7$  Gaussian filter with standard deviation of 1.6, and then downsampling the blurred image by a scaling factor of 3 in both horizontal and vertical directions. The additive Gaussian noise with standard deviation of 5 is also added to the LR images, making this IR problem more challenging. Since human visual system is more sensitive to luminance changes, we only apply the IR methods to the luminance component and use the simple bicubic interpolator for the chromatic components.

We compare the proposed CSR based IR method with some recently developed image super-resolution methods, i.e., the softcut method of [12], the TV-based method of [22] and the sparse representation based method of [27]. Since the sparsity-based method in [27] cannot perform the resolution upscaling and deblurring simultaneously, as suggested by [27] we apply the iterative back-projection to the output of the method [27] to remove the blur.

The PSNR results of the competing methods on a set of 9 natural images are reported in Table 2. From Table 2, we can conclude that the proposed CSR approach significantly outperforms all the other competing methods. This demonstrates the superiority of the CSR model for solving image

inverse problems. Some subjective comparison between the CSR and other methods are shown in Fig. 4. We can see that the TV-based method [22] tends to generate piecewise constant structures; the softcut method [12] produces over-smoothed image local structures; the image edges reconstructed by the sparsity-based method [27] contain some visible artifacts. Obviously, the image reconstructed by CSR gives the best visual quality. The reconstructed edges are much sharper than all the other three competing methods, and more image fine structures are recovered.

## 5. Conclusion

Image restoration (IR) is a fundamental topic in image processing and computer vision applications, and it has been widely studied. In this paper, we investigated IR with the sparse coding techniques. To better understand the effectiveness of sparse coding for IR, we introduced the concept of sparse coding noise (SCN), and it was empirically found that SCN follows Laplacian distributions. To suppress SCN and thus improve the quality of IR, the centralized sparse representation (CSR) model was proposed by exploiting the image nonlocal self-similarity. In addition to the local sparsity, we also enforced the sparse coefficients to have small SCN, i.e., to be close to their distribution cen-

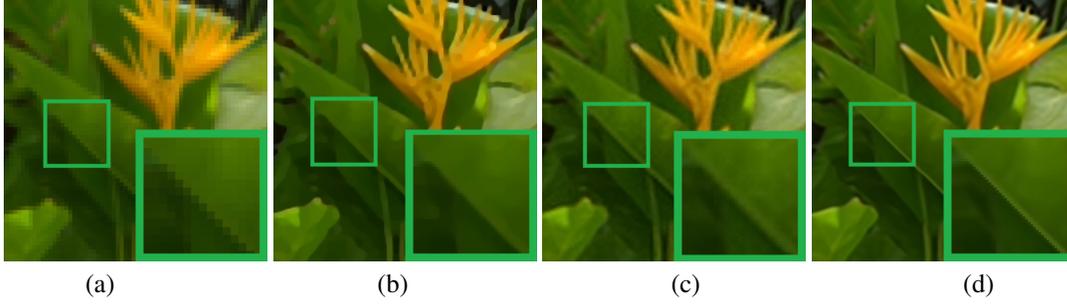


Figure 4. Super-resolution performance comparison for *Plant* image, (scaling factor 3,  $\sigma_n = 0$ ). (a) LR image; (b) TV-based method [22] (PSNR=31.34dB); (c) Sparsity-based method [27] (PSNR=31.55dB); (d) the proposed CSR method (PSNR=34.00 dB).

Table 1. PSNR(dB) results of the deblurred images.

9×9 uniform blur, $\sigma_n = \sqrt{2}$										
Images	<i>Butterfly</i>	<i>Boats</i>	<i>C. Man</i>	<i>Starfish</i>	<i>Parrot</i>	<i>Lena</i>	<i>Barbara</i>	<i>Peppers</i>	<i>Leaves</i>	<i>Average</i>
FISTA [2]	28.37	29.04	26.82	27.75	29.11	28.33	25.75	28.43	26.49	27.79
SA-DCT [18]	27.50	29.25	27.02	28.11	30.07	28.43	26.58	28.36	27.04	28.04
BM3D [11]	27.21	29.93	27.30	28.61	30.50	28.97	<b>27.99</b>	28.60	27.45	28.51
CSR	<b>29.75</b>	<b>31.10</b>	<b>28.55</b>	<b>30.30</b>	<b>32.09</b>	<b>29.95</b>	27.93	<b>29.64</b>	<b>29.97</b>	<b>29.95</b>
9×9 uniform blur, $\sigma_n = 2$										
FISTA [2]	27.73	27.93	26.13	27.50	28.88	27.40	25.24	27.42	26.03	27.14
SA-DCT [18]	26.46	28.14	26.11	27.14	29.10	27.58	25.75	27.57	25.86	27.08
BM3D [11]	26.56	29.21	26.61	27.97	29.75	28.34	<b>27.26</b>	28.02	26.60	27.81
CSR	<b>28.66</b>	<b>30.48</b>	<b>27.68</b>	<b>29.00</b>	<b>30.57</b>	<b>29.23</b>	27.20	<b>29.05</b>	<b>28.64</b>	<b>28.94</b>
Gaussian blur with standard deviation 1.6, $\sigma_n = \sqrt{2}$										
FISTA [2]	30.36	29.36	26.80	29.65	31.23	29.47	25.03	29.42	29.33	28.97
SA-DCT [18]	29.85	30.28	27.44	30.84	32.46	30.43	27.00	29.22	29.70	29.69
BM3D [11]	29.01	30.63	27.46	30.71	32.22	30.69	<b>28.19</b>	29.03	29.67	29.74
CSR	<b>30.75</b>	<b>31.40</b>	<b>28.24</b>	<b>32.31</b>	<b>33.44</b>	<b>31.23</b>	27.81	<b>30.17</b>	<b>31.44</b>	<b>30.70</b>
Gaussian blur with standard deviation 1.6, $\sigma_n = 2$										
FISTA [2]	29.67	27.89	25.94	29.18	30.74	28.00	24.54	27.94	28.62	28.07
SA-DCT [18]	29.42	29.88	26.99	30.04	31.79	29.96	26.08	28.90	29.16	29.13
BM3D [11]	28.56	30.21	27.08	30.23	31.72	30.28	<b>27.02</b>	28.73	29.10	29.21
CSR	<b>30.14</b>	<b>31.19</b>	<b>27.81</b>	<b>31.47</b>	<b>32.60</b>	<b>30.94</b>	26.53	<b>30.03</b>	<b>30.56</b>	<b>30.09</b>

ters. The local and nonlocal redundancy induced sparsity, which are both characterized by the  $l_1$ -norm, are unified into a variational formulation. A Bayesian interpretation of the CSR model was then provided to accurately determine the regularization parameters. Experimental results on IR demonstrated that the CSR image restoration approach can significantly outperform other leading IR methods.

## 6. Acknowledgements

This work was partially supported by grant NSFC (No. 61033004, 60736043, 61070138, and 61072104), the Fundamental Research Funds of the Central Universities of China (No. K50510020003), and HK RGC General Re-

search Fund (PolyU 5375/09E).

## References

- [1] M. Aharon, M. Elad, and A. Bruckstein. K-svd: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.*, 54(11):4311–4322, Nov. 2006. **1, 4**
- [2] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. On Image Process.*, 18(11):2419–2434, Nov. 2009. **5, 6, 7**
- [3] D. Bertsekas, A. Nedic, and A. Ozdaglar. *Convex Analysis and Optimization*. Athena Scientific, 2003. **5**

Table 2. PSNR(dB) results(luminance components) of the reconstructed HR images.

Noiseless										
Images	<i>Butterfly</i>	<i>flower</i>	<i>Girl</i>	<i>Pathenon</i>	<i>Parrot</i>	<i>Raccoon</i>	<i>Bike</i>	<i>Hat</i>	<i>Plants</i>	<i>Average</i>
TV [22]	26.56	27.51	31.24	26.00	27.85	27.54	23.66	29.20	31.34	27.88
Softcut [12]	24.74	27.31	31.82	25.95	27.99	27.82	23.15	29.50	31.19	27.72
Sparsity [27]	24.70	27.87	32.87	26.27	28.70	28.51	23.23	29.63	31.55	28.15
Proposed CSR	<b>28.19</b>	<b>29.54</b>	<b>33.68</b>	<b>27.23</b>	<b>30.68</b>	<b>29.29</b>	<b>24.72</b>	<b>31.33</b>	<b>34.00</b>	<b>29.85</b>
Noisy										
TV [22]	25.49	26.57	29.86	25.35	27.01	26.74	23.11	28.13	29.70	26.88
Softcut [12]	24.53	26.98	31.30	25.72	27.69	27.48	22.91	29.13	30.57	27.37
Sparsity [27]	23.61	26.60	30.71	25.40	27.15	27.22	22.45	28.31	29.57	26.78
Proposed CSR	<b>26.84</b>	<b>28.07</b>	<b>32.03</b>	<b>26.40</b>	<b>29.47</b>	<b>28.03</b>	<b>23.78</b>	<b>29.96</b>	<b>31.73</b>	<b>28.48</b>

- [4] J. Biemond, R. Lagendijk, and R. Mersereau. Iterative methods for image deblurring. *Proceedings of the IEEE*, 78(5):856–883, 1990. **1**
- [5] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, Feb. 2009. **1**
- [6] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. In *IEEE Int. Conf. on CVPR*, volume 2, pages 60–65, 2005. **1, 2, 3**
- [7] E. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier Analysis and Applications*, 14:877–905, Dec. 2008. **4**
- [8] P. Chatterjee and P. Milanfar. Clustering-based denoising with locally learned dictionaries. *IEEE Trans. Image Processing*, 18(7):1438–1451, July 2009. **5**
- [9] S. Chen, D. Donoho, and M. Saunders. Atomic decompositions by basis pursuit. *SIAM Review*, 43:129–159, 2001. **1, 2**
- [10] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Trans. Image Processing*, 16(8):2080–2095, Aug. 2007. **2, 5**
- [11] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image restoration by sparse 3d transform-domain collaborative filtering. In *SPIE Conference Series*, 2008. **5, 6, 7**
- [12] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, and A. K. Katsaggelos. Softcuts: a soft edge smoothness prior for color image super-resolution. *IEEE Trans. Image Processing*, 18(5):969–981, May 2009. **1, 6, 8**
- [13] I. Daubechies, M. Defriese, and C. DeMol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.*, 57:1413–1457, 2004. **1, 2, 5**
- [14] W. Dong, G. Shi, L. Zhang, and X. Wu. Super-resolution with nonlocal regularized sparse representation. In *Proc. of SPIE Vis. Comm. and Image Process. (VCIP)*, volume 7744, pages 77440H–1–77440H–10, 2010. **1, 5**
- [15] W. Dong, L. Zhang, G. Shi, and X. Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Trans. Image Processing*, 20(7):1838–1857, July 2011. **1, 5**
- [16] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.*, 15(12):3736–3745, Dec. 2006. **1, 2**
- [17] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camerashake from a single image. *ACM Trans. Graph. (SIGGRAPH)*, pages 787–794, 2006. **5, 6**
- [18] A. Foi, V. Katkovnik, and K. Egiazarian. Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images. *IEEE Trans. Image Process.*, 16(5):1395–1411, May 2007. **5, 7**
- [19] J. Mairal, F. Bach, J. Ponce, and G. Sapiro. Online dictionary learning for sparse coding. In *Proc. of the 26th Annual Int. Conf. on Machine Learning*, pages 689–696, 2009. **1**
- [20] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *Proc. of the IEEE ICCV*, Tokyo, Japan, 2009. **1, 2, 5**
- [21] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Trans. on Image Processing*, 17(1):53–69, Jan. 2008. **1**
- [22] A. Marquina and S. J. Osher. Image super-resolution by tv-regularization and bregman iteration. *J. Sci. Comput.*, 37:367–382, 2008. **6, 7, 8**
- [23] B. Olshausen and D. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996. **1**
- [24] B. Olshausen and D. Field. Sparse coding with an overcomplete basis set: a strategy employed by v1? *Vision Research*, 37:3311–3325, 1997. **1**
- [25] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992. **1**
- [26] L. Sendur and I. W. Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Trans. Signal Process.*, 50(11):2744–2756, Nov. 2002. **4**
- [27] J. Yang, J. Wright, Y. Ma, and T. Huang. Image super-resolution as sparse representation of raw image patches. In *Proc. of the IEEE CVPR*, Jun. 2008. **1, 6, 7, 8**
- [28] L. Zhang, W. Dong, D. Zhang, and G. Shi. Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, 43:1531–1549, Apr. 2010. **5**