Journal of Information Hiding and Multimedia Signal Processing ©2014 ISSN 2073-4212
Ubiquitous International Volume **5**, Number **4**, October **2014**

A Sparse Representation Based Super-resolution Image Reconstruction Scheme Utilizing Dual Dictionaries

Hao-Xian Wang

School of Information and Electrical Engineering
Harbin Institute of Technology at Weihai
Weihai 264209, P. R. China
haoxianwang@hitwh.edu.cn

Zhe-Ming Lu

School of Aeronautics and Astronautics
Zhejiang University
Hangzhou, 310027, P.R. China
zheminglu@zju.edu.cn

Yong Zhang

School of Information and Electrical Engineering
Harbin Institute of Technology at Weihai
Weihai 264209, P. R. China.
zhangyong@hitwh.edu.cn

Received September 2013; revised April 2014
(Communicated by Hao-Xian Wang)

Abstract. *Super-resolution (SR) image reconstruction is a technique to generate a high resolution (HR) image from several low resolution (LR) images of the same scene, which can improve the visual effect of images or serve as a pre-processing technique. Among various SR image reconstruction schemes, sparse representation based SR image reconstruction schemes have become the current research focus because of their excellent reconstruction quality. In this paper, an SR image reconstruction algorithm based on sparse representation with a redundant dictionary is proposed. In our algorithm, the redundant dictionary and the coding dictionary are trained jointly, and the representation coefficients can be calculated by simply multiplying the input signal by the coding dictionary, which can reduce the computational complexity greatly. The proposed algorithm makes full use of some constraint terms such as the consistence of sparse representation coefficients between the HR image and corresponding LR images, a sparsity prior, autoregressive models and nonlocal means regularization to build up the cost function for SR image reconstruction, and then solves this function based on the iterative shrinkage method to obtain the target HR image. Experimental results demonstrate that the proposed method can achieve great improvement in terms of visual effect, PSNR and SSIM.*
Keywords: Super-resolution image reconstruction, Sparse representation, Iterative shrinkage, Nonlocal means, Redundant dictionary.

1. **Introduction.** Super resolution (SR) image reconstruction is a technique to restore a high resolution (HR) image from a single or multiple low resolution (LR) images, which can be used to display LR images obtained from various LR imaging devices (such as

mobile telephones and surveillance devices) on HR devices (e.g., HDTV). SR image reconstruction techniques have been widely applied in various fields including medical imaging, satellite remote sensing, military reconnaissance, and city security systems. With many years development, SR image reconstruction techniques can be classified into two main categories according to different types of input information, i.e., conventional multi-frame based methods [1-3], and single-frame based algorithms [4-7].

Conventional multi-frame based SR image reconstruction algorithms make full use of multiple LR images captured from the same scene to generate a SR image. Three typical traditional methods that have been widely researched are maximum a posteriori (MAP) [8], projection onto a convex set (POCS) [9] and iterative back projection (IBP) [10]. They are based on the same model, and they are designed to obtain the HR images by imposing different additional constraints on ill-posed problems. In order to take advantage of the information from multiple input LR images, it is necessary to perform sub-pixel motion vector estimation on the sequence of input images. The main idea of this kind of schemes is to generate the model based on the LR images and then solve this model conversely to obtain the SR image. However, too many parameters such as motion vectors and the settings of fuzzy matrices for these conventional schemes are required to be estimated, while the input information is insufficient, the reconstruction results are therefore unsatisfactory.

Single-frame based schemes perform SR image reconstruction only based on one input LR image. In general, single-frame based schemes can be classified into two categories. The first category embodies interpolation-based algorithms that are relatively simple and real-time, and they often serve as the preprocessing step for other SR image reconstruction schemes. Two typical ones are bilinear interpolation and bicubic interpolation [4]. They do not consider the intrinsic structure such as edges in images, resulting in image blurring or false edges. Sun et al. proposed [11] to perform the interpolation-based image reconstruction first, and then adjust the edges of the reconstructed image according to a priori HR edge information obtained from a large number of natural HR images. However, this method is apt to cause distortions after reconstruction. The second category embodies learning-based schemes, which first learn a prior information based on the LR images and their corresponding HR images in the database, and then guide the SR image reconstruction based on the obtained a prior information. Freeman et al. [5] modeled the spatial relationship between LR patches and HR patches using a Markov network, and then estimated the missing high-frequency content due to degradation, and finally added it to the initial interpolation result to obtain the output HR image. Reference [12] proposed an improved scheme by introducing the so-called Primal Sketch Priors to the constraints during the reconstruction process. In general, the algorithms in this category can overcome the limitation in the number of times of resolution improvement compared with other SR image reconstruction schemes, however, since they need many paired low-resolution and high-resolution image blocks, their computational complexity is very high.

With the development of the techniques for solving the problem of l_1 norm, there emerges a new branch of learning-based schemes, namely, the single-frame SR image reconstruction based on sparse representation. Natural images are sparse in some domains, namely, they can be linearly represented by several sparse bases, and the representation coefficients are sparse too (i.e., most of the coefficients are equal to zero or very close to zero). Since the sparse representation is universal for images, it has been successfully applied to many fields such as image encoding, image restoration and image denoising [13-16]. References [7] and [16] presented a SR image reconstruction scheme based on sparse representation. This scheme first obtains a pair of redundant dictionaries trained

from an image database, assuming that the sparse coefficients of the LR image patches decomposed using the LR redundant dictionary are the same as those of the corresponding HR image patches decomposed using the HR redundant dictionary. During the image reconstruction process, the LR image patches are first represented using the LR redundant dictionary to obtain the sparse coefficients, then the initially estimated SR image is obtained by combining the SR redundant dictionary, and finally the global constraints are considered to obtain the optimized results. This scheme can achieve a very good effect in normal cases, however it cannot achieve good effect in the case that the input image is much blurred or the number of times enlarged is big. Reference [17] utilized the PCA (Principal Component Analysis) technique to train multiple sub-dictionaries, and ARM (Autoregressive Models) and NL-M (NonLocal Means) constraints are added in the reconstruction equation to constrain the target HR image. To solve the reconstruction equation, the IS (Iterative Shrinkage) based method is used, and thus a good reconstruction effect can be obtained.

For sparse representation based image reconstruction schemes, one of main steps is to construct and train the redundant dictionaries. Reference [14] adopted the K-SVD algorithm to train the redundant dictionaries, however the computational complexity is very high and the training performance is not so good in the case that there is large similarity between samples. Reference [17] utilized the PCA method to train multiple orthogonal sub-dictionaries, and then represented the input image based on sparse representation. However, the PCA orthogonal dictionaries are only effective for the Euclidean structure but ineffective for nonlinear structures. Fortunately, since the l_1 norm is adopted, the redundant dictionaries can be used to overcome this shortcoming to some degree. Reference [18] proposed a very efficient solution to redundant dictionaries, and the obtained dictionaries can be used to achieve a very good sparse decomposition, however, the computational complexity is very high. Reference [19] obtained the so-called encoding dictionary at the same time of learning the redundant dictionaries. Since the sparse coefficients can be obtained by multiplying the encoding dictionary by the input signal, the computational complexity can be greatly reduced, however, the accuracy of sparse coefficients is not high. Because of above situation, this paper presents to solve the sparse coefficients from the point of view of super completed redundant dictionaries. The redundant dictionaries are first learned using the dual dictionary method in [19], and then the iteration is divided into two parts, the first part is to solve the sparse coefficients based on the encoding dictionary, while the second part is to solve the sparse coefficients based on the feature sign search algorithm in [18]. The remainder of this paper is organized as follows. Section 2 introduces a priori constraints of the SR image reconstruction problem based on sparse representation. Section 3 presents our dictionary training scheme. Section 4 proposes our iterative shrinkage based sparse representation algorithm. Finally, various experimental results in Section 5 demonstrate the efficiency of our algorithm.

2. **A Priori Constraints.**

Super resolution image reconstruction is actually a process to select the constraint components according to the input image and compose an equation and then solve it, where the component selection concerns the reconstruction effect. With the development of digital image processing technology, people have proposed many a prior constraints. In order to achieve a satisfactory reconstruction effect, this paper performs a prior constraints based on the following idea:

(1) The LR image formation model: for single-frame SR image reconstruction, the LR image $\boldsymbol{Y}$ is obtained by blurring an ideal SR image $\boldsymbol{X}$ with a point spread function $\boldsymbol{H}$ and down sampling it and imposing noise on it, the corresponding mathematical formula

can be expressed as follows:

$$\boldsymbol{Y} = \boldsymbol{DHX} + \boldsymbol{v} \tag{1}$$

where $\boldsymbol{D}$ is the down-sampling matrix and $\boldsymbol{v}$ denotes the noise. In fact, SR image reconstruction is the process of solving the inverse problem based on the input LR image $\boldsymbol{Y}$ to get the SR image $\boldsymbol{X}$. Because the number of unknown quantities is more than the number of known quantities, this problem is an ill-posed problem, which requires some other a priori information.

(2) Sparse a prior: natural images are generally compressible, that is, they can be sparse coded [20] in a certain domain. With the development of l_1 norm problem solving techniques, sparse representation is more and more widely applied in digital image processing for reverse problem solving [21]. For the input $\boldsymbol{x} \in R^n$, based on the given redundant dictionary $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \cdots, \boldsymbol{\varphi}_m] \in R^{n \times m}$, assume the sparse coded coefficient vector is $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_m]^T$ (most of the coefficients are close to 0), we have:

$$\boldsymbol{x} \approx \boldsymbol{\Phi\alpha} \tag{2}$$

Because the l_0 norm problem is not a convex problem but a NP problem, we adopt the l_1 norm to solve the sparse coefficients as follows:

$$\boldsymbol{\alpha} = arg \min_{\boldsymbol{\alpha}} \left\{ \|\boldsymbol{x} - \boldsymbol{\Phi\alpha}\|_2^2 + \lambda \left|\boldsymbol{\alpha}\right|_1 \right\} \tag{3}$$

where λ is a constant. To solve this problem, this paper adopts the feature sign search algorithm [18].

3. **Redundant Dictionary Training.** The sparse dictionary $\boldsymbol{\Phi}$ and ARM weighting parameters, especially the sparse dictionary, play a key role in SR image reconstruction. This paper adopts the dual dictionary learning algorithm for redundant dictionary training. After the training process, we can obtain two dictionaries, one is the redundant dictionary and the other is the encoding dictionary. The redundant dictionary training process includes two main steps, one is to establish the training samples and the other is to solve the dictionary:

(1) Establishment of training samples

In order to train the redundant dictionary $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2, \cdots, \boldsymbol{\varphi}_m] \in R^{n \times m}$, samples of size $\sqrt{n} \times \sqrt{n}$ are required to be cropped from SR images in the database. Each SR image in the database is divided into blocks of size $\sqrt{n} \times \sqrt{n}$, then all the obtained image blocks are arranged into a column vector, denoted by $\boldsymbol{s}_i$ ($\boldsymbol{s}_i \in R^n$). Through computing the variance of $\boldsymbol{s}_i$ (denoted by $\mathrm{Var}(\boldsymbol{s}_i)$), remaining the column vectors whose $\mathrm{Var}(\boldsymbol{s}_i)$ is greater than a certain threshold, and removing the column vectors that are relatively smooth, we can finally obtain a set of training samples $\boldsymbol{S} = \{\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_M\}$, where M denotes the size of the set of samples.

(2) Training of dual dictionaries

The dual dictionary training method can obtain the linear mapping of the redundant dictionary, i.e., the coding dictionary, at the same time of training the redundant dictionary. In general, repeated iterations are required for solving the sparse coefficients of a signal, and thus the computational complexity is relatively high. In fact, the sparse coefficients can be obtained through directly multiplying the coding dictionary by the input signal, which can greatly reduce the computational complexity, however, the obtained sparse coefficients are not accurate and the error is large. Compromising above two methods, this paper proposes a new solution, i.e., the iterative shrinkage process is divided into two parts, in the first part the sparse coefficients are obtained by multiplying

the coding dictionary, in the second part the feature sign search algorithm is adopted to solve the sparse coefficients, which greatly improves the reconstruction effect.

The equation for solving the redundant dictionary can be described as:

$$\{\boldsymbol{\Phi}, \boldsymbol{\Theta}\} = arg \min_{\boldsymbol{D},\boldsymbol{\Theta}} \left\{ \|\boldsymbol{S} - \boldsymbol{\Phi}\boldsymbol{\Theta}\|_2^2 + \lambda \left|\boldsymbol{\Theta}\right|_1 \right\} \tag{4}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_M] \in R^{m \times M}$ denotes the set of sparse coefficients for all samples in the sample set $\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_M]$. The equation for dual dictionary training can be described as:

$$\begin{aligned} \{\boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{\Theta}\} = \min_{\boldsymbol{\Phi},\boldsymbol{\Psi},\boldsymbol{\Theta}} \left\{ \|\boldsymbol{S} - \boldsymbol{\Phi}\boldsymbol{\Theta}\|_2^2 + \eta \|\boldsymbol{\Theta} - \boldsymbol{\Psi}\boldsymbol{S}\|_2^2 + \lambda \left|\boldsymbol{\Theta}\right|_1 \right\} \\ s.t. \, \|\boldsymbol{\varphi_i}\|_2^2 \leq 1, \|\boldsymbol{\psi_i}\|_2^2 \leq 1 \end{aligned} \tag{5}$$

where $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \cdots, \boldsymbol{\psi}_M] \in R^{m \times n}$ is the coding dictionary,$\boldsymbol{\psi}_i$ is the i-th column of $\boldsymbol{\Psi}$, and η is a constant. This equation can be solved by using the iterative method as follows.

1) First, $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are fixed, and $\boldsymbol{\Theta}$ is updated by transforming Eq. (5) as follows:

$$\boldsymbol{\Theta} = arg \min_{\boldsymbol{\Theta}} \left\{ \|\boldsymbol{S} - \boldsymbol{\Phi}\boldsymbol{\Theta}\|_2^2 + \eta \|\boldsymbol{\Theta} - \boldsymbol{\Psi}\boldsymbol{S}\|_2^2 + \lambda \left|\boldsymbol{\Theta}\right|_1 \right\} \tag{6}$$

Then we can have the following iterative expression for $\boldsymbol{\Theta}$:

$$\boldsymbol{\Theta}^{k+1} = T_{\lambda/2\sigma_\Theta} \left[\left(1 - \frac{\eta}{\sigma_\Theta} \right) \boldsymbol{\Theta}^k + \frac{1}{\sigma_\Theta} \left(\boldsymbol{\Phi}^T \left(\boldsymbol{S} - \boldsymbol{\Phi}\boldsymbol{\Theta}^k \right) + \eta \boldsymbol{\Psi}\boldsymbol{S} \right) \right] \tag{7}$$

Where $\sigma_\Theta = 2 \left\| \boldsymbol{\Phi}^{\boldsymbol{T}}\boldsymbol{\Phi} + \eta \boldsymbol{I} \right\|_F$,$\boldsymbol{I}$ is the unit matrix. If T_τ [] is defined as the thresholding operation, and $\boldsymbol{X}_{i,j}$ denotes the element of the matrix $\boldsymbol{X}$ at the location (i, j), we have:

$$\left(T_\tau \left[\boldsymbol{X}\right]\right)_{i,j} = sign\left(\boldsymbol{X}_{i,j}\right) max\left\{ \left|\boldsymbol{X}_{i,j}\right| - \tau, 0 \right\} \tag{8}$$

2) Second, $\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ are fixed,and $\boldsymbol{\Phi}$ can be solved by transforming Eq. (5) as follows:

$$\boldsymbol{\Phi} = arg \min_{\boldsymbol{\Phi}} \|\boldsymbol{S} - \boldsymbol{\Phi}\boldsymbol{\Theta}\|_2^2 \quad s.t. \quad \|\boldsymbol{\varphi_i}\|_2^2 \leq 1 \tag{9}$$

If we define the operation $\pi\left(\boldsymbol{d}\right) = \boldsymbol{d}/max\left\{1, \|\boldsymbol{d}\|\right\}$ which is used to project a vector to its unit length, then the updating expression of $\boldsymbol{\Phi}$ can be described as follows:

$$\boldsymbol{\Phi}^{k+1} = \pi_\Phi \left(\boldsymbol{\Phi}^k + \frac{1}{\sigma_\Phi} \left(\boldsymbol{S} - \boldsymbol{\Phi}^k \boldsymbol{\Theta} \right) \boldsymbol{\Theta}^T \right) \tag{10}$$

where π_Φ denotes the operation of projecting $\boldsymbol{\Phi}$ to its unit length by the operator $\pi\left(\right)$, and $\sigma_\Phi = 2 \left\| \boldsymbol{\Theta}\boldsymbol{\Theta}^T \right\|_F$.

3) Finally, $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ are fixed, and $\boldsymbol{\Psi}$ can be solved by transforming Eq. (5) as follows:

$$\boldsymbol{\Psi} = arg \min_{\Psi} \eta \|\boldsymbol{\Theta} - \boldsymbol{\Psi}\boldsymbol{S}\|_2^2 \quad s.t. \quad \|\boldsymbol{\psi_i}\|_2^2 \leq 1 \tag{11}$$

Similar to $\boldsymbol{\Phi}$, the updating expression of $\boldsymbol{\Psi}$ can be described as follows:

$$\boldsymbol{\Psi}^{k+1} = \pi_\Psi \left(\boldsymbol{\Psi}^k + \frac{1}{\sigma_\Psi} \left(\boldsymbol{\Theta} - \boldsymbol{\Psi}^k \boldsymbol{S} \right) \boldsymbol{S}^T \right) \tag{12}$$

where π_Ψ denotes the operation of projecting $\boldsymbol{\Psi}$ to its unit length by the operator $\pi\left(\right)$, and $\sigma_\Psi = 2 \left\| \boldsymbol{S}\boldsymbol{S}^T \right\|_F$.

4. **Iterative Shrinkage Based Sparse Representation Algorithm.** After training the redundant dictionaries, we can perform the sparse representation based on the redundant dictionaries. For each image block, we adaptively select the ARM weighting parameters to constrain its center pixel, and perform the non-local mean constraints on it, then utilize the iterative shrinkage method to solve the equation, obtaining the final SR image output.

Let $\boldsymbol{x}_i \in R^n$ denote the image block of size $\sqrt{n} \times \sqrt{n}$ obtained from the SR image $\boldsymbol{X}$, we can define $\boldsymbol{x}_i = \boldsymbol{R_i X}$, $i = 1, 2, \cdots, N$, where $\boldsymbol{R}_i$ is the cropping matrix used to obtain $\boldsymbol{x}_i$ from $\boldsymbol{X}$. According to the sparse prior, $\boldsymbol{x}_i$ can be represented by using the redundant dictionary, namely,$\boldsymbol{x}_i = \boldsymbol{\Phi \alpha}_i$, thus $\boldsymbol{X}$can be expressed as [17]:

$$\mathbf{X} = \left(\sum_{i=1}^{N} \boldsymbol{R}_i^T \boldsymbol{R}_i \right)^{-1} \sum_{i=1}^{N} \boldsymbol{R}_i^T \boldsymbol{\Phi \alpha}_i \tag{13}$$

For convenience, we define:

$$\mathbf{X} = \boldsymbol{\Phi \cdot \Theta} = \left(\sum_{i=1}^{N} \boldsymbol{R}_i^T \boldsymbol{R}_i \right)^{-1} \sum_{i=1}^{N} \boldsymbol{R}_i^T \boldsymbol{\Phi \alpha}_i \tag{14}$$

where $\boldsymbol{\Theta} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha_N}]$. According to Eqs. (1), (3) and (14), the final SR reconstruction equation can be expressed as follows:

$$\boldsymbol{\Theta} = arg \min_{\boldsymbol{\Theta}} \left\{ \| \boldsymbol{Y - DH\Phi \cdot \Theta} \|_2^2 + \lambda \left| \boldsymbol{\Theta} \right|_1 \right\} \tag{15}$$

After obtaining $\boldsymbol{\Theta}$, we can get the SR image $\boldsymbol{X}$according to Eq. (14).

4.1. **Adaptive Selection of ARM Weights.** The main idea of the autoregressive model is to make use of the pixels around the current pixel to constrain the current pixel value and to minimize the error between the current pixel value and the weighted sum of the neighboring pixels. The training of weighting parameters for the autoregressive model is a process to get the weighting parameters of the surrounding pixels of the current pixel. This paper uses the method in [17] to train the weighting parameters, obtaining the ARM weighting parameters $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_K\}$ and corresponding index vectors $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \cdots, \boldsymbol{\mu}_K\}$.

Since different images have different variances of image blocks as well as different ARM weights, for each input image vector $\boldsymbol{x}_i$, we need to find the most suitable ARM parameter vector to constrain the center pixel of the input image vector $\boldsymbol{x}_i$. Similar to the training process of the ARM weighting parameters, first of all we need to obtain the high frequency information vector $\boldsymbol{x}_i^h$ of $\boldsymbol{x}_i$. The selection of ARM parameter vectors can be expressed as follows:

$$k_i = arg \min_{k} \left\| \boldsymbol{x}_i^h - \boldsymbol{\mu}_k \right\|_2 \tag{16}$$

Thus the obtained ARM weighting parameters for $\boldsymbol{x}_i$ should be $\boldsymbol{a}_{k_i}$. Assume x_i is the center pixel of $\boldsymbol{x}_i$, and $\boldsymbol{\chi}_i$ is the vector consisting of the neighboring pixels adjacent to the center pixel x_i, then the optimal current pixel value x_i should minimize $\left\| x_i - \boldsymbol{a}_{k_i}^T \boldsymbol{\chi}_i \right\|_2^2$. Adding this constraint to Eq. (15), the reconstruction equation can be rewritten as:

$$\boldsymbol{\Theta} = arg \min_{\boldsymbol{\Theta}} \left\{ \| \boldsymbol{Y - DH\Phi \cdot \Theta} \|_2^2 + \lambda \left| \boldsymbol{\Theta} \right|_1 + \gamma \sum_i \left\| x_i - \boldsymbol{a}_{k_i}^T \boldsymbol{\chi}_i \right\|_2^2 \right\} \tag{17}$$

Where γ is a constant balancing factor. Here, for convenience, $\sum_i \left\|x_i - \boldsymbol{a}_{k_i}^T \boldsymbol{\chi}_i\right\|_2^2$, where $\boldsymbol{I}$ is the unit matrix, and the matrix $\boldsymbol{A}$ satisfies:

$$\boldsymbol{A}(i,j) = \begin{cases} a_i & if\ x_j\ belongs\ to\ \boldsymbol{\chi}_i, a_i \in \boldsymbol{a}_{k_i} \\ 0 & otherwise \end{cases} \tag{18}$$

At this time, the reconstruction equation can be improved as follows:

$$\boldsymbol{\Theta} = arg\min_{\boldsymbol{\Theta}} \left\{ \|\boldsymbol{Y} - \boldsymbol{DH\Phi} \cdot \boldsymbol{\Theta}\|_2^2 + \lambda |\boldsymbol{\Theta}|_1 + \gamma \|(\boldsymbol{I} - \boldsymbol{A})\, \boldsymbol{\Phi} \cdot \boldsymbol{\Theta}\|_2^2 \right\} \tag{19}$$

4.2. **Non-local Mean Filtering.** The principle of non-local mean filtering is very simple, that is, for each current image block, we can find several similar blocks to constrain it. For each image block $\boldsymbol{x}_i$, we find all image blocks similar to it in the whole image. Assume $\boldsymbol{x}_i^s$ is one of blocks similar to $\boldsymbol{x}_i$, then $e_i^s = \|\boldsymbol{x}_i^s - \boldsymbol{x_i}\|_2^2$ should be small enough. If we calculate the weighted sum of all the center pixels in these similar blocks, the results should satisfy [17] $x_i \approx \sum_{s=1}^{L} b_i^s x_i^s$, where b_i^s denotes the weight allocated to $\boldsymbol{x}_i^s$. Obviously, the more two blocks are similar, the greater weight should be allocated. Thus, we can describe the expression for weight calculation as $b_i^s = exp\left(-e_i^s/h\right)/c_i$, where c_i is the normalization factor,h is a constant,x_i^s is the center pixel of $\boldsymbol{x}_i^s$, and L is the number of similar blocks (in this paper, we set $L = 7$). Assume $\boldsymbol{b}_i = \left[b_i^1, b_i^2, \cdots, b_i^L\right]$ is the weighting vector,$\boldsymbol{\beta}_i = \left[x_i^1, x_i^2, \cdots, x_i^L\right]$ is the set of center pixels, we should minimize $\sum_i \left\|x_i - \boldsymbol{b}_i^T \boldsymbol{\beta}_i\right\|_2^2$ for the whole image. Similarly, we use $\|(\boldsymbol{I} - \boldsymbol{B})\, \boldsymbol{X}\|_2^2$ to describe $\sum_i \left\|x_i - \boldsymbol{b}_i^T \boldsymbol{\beta}_i\right\|_2^2$, where $\boldsymbol{B}$ should satisfy:

$$\boldsymbol{B}(i,s) = \begin{cases} b_i^s & if\ x_i^s\ belongs\ to\ \boldsymbol{\beta}_i, b_i^s \in \boldsymbol{b}_i \\ 0 & otherwise \end{cases} \tag{20}$$

Thus, the reconstruction equation can be rewritten as:

$$\boldsymbol{\Theta} = arg\min_{\boldsymbol{\Theta}} \left\{ \|\boldsymbol{Y} - \boldsymbol{DH\Phi} \cdot \boldsymbol{\Theta}\|_2^2 + \lambda |\boldsymbol{\Theta}|_1 + \gamma \|(\boldsymbol{I} - \boldsymbol{A})\, \boldsymbol{\Phi} \cdot \boldsymbol{\Theta}\|_2^2 + \eta \|(\boldsymbol{I} - \boldsymbol{B})\, \boldsymbol{\Phi} \cdot \boldsymbol{\Theta}\|_2^2 \right\} \tag{21}$$

The above equation can be solved by using the iterative shrinkage algorithm, which can be illustrated as follows:

Step 1: Initialization.

Step 1.1: Use the bi-cubic interpolation scheme to expand the LR images to the required sizes, as the initial input $\boldsymbol{X}^{(0)}$.

Step 1.2: For each block in the obtained SR image, select proper ARM weighting parameters, and calculate the similar block, obtaining the set of weights $\boldsymbol{b}_i$.

Step 1.3: Initialize matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ according to the calculation results in Step 1.2.

Step 1.4: Preset the parameters γ, η, P, e Mid_Iter and the maximum number of iterations Max_Iter, and set $k = 0$.

Step 2: Iterative Process: if k is not greater than Max_Iter and $\left\|\boldsymbol{X}^{(k+1)} - \boldsymbol{X}^{(k)}\right\|_2^2 / \left\|\boldsymbol{X}^{(k)}\right\|_2^2 > e$, perform the following sub-steps; Otherwise, stop the iteration process, go to Step 3.

Step 2.1: Calculate

$\boldsymbol{X}^{(k+1/2)} = \boldsymbol{X}^{(k)} + (\boldsymbol{DH})^T\, \boldsymbol{Y} - (\boldsymbol{DH})^T\, \boldsymbol{DHX}^{(k)} - \gamma\, (\boldsymbol{I} - \boldsymbol{A})^T\, (\boldsymbol{I} - \boldsymbol{A})\, \boldsymbol{X}^{(k)}$
$- \eta\, (\boldsymbol{I} - \boldsymbol{B})^T\, (\boldsymbol{I} - \boldsymbol{B})\, \boldsymbol{X}^{(k)}$

Step 2.2: If k is less than Mid_Iter, then

$\boldsymbol{\Theta}^{(k+1/2)} = \left[\boldsymbol{\Psi}\boldsymbol{R}_1\boldsymbol{X}^{(k+1/2)}, \boldsymbol{\Psi}\boldsymbol{R}_2\boldsymbol{X}^{(k+1/2)} \cdots \boldsymbol{\Psi}\boldsymbol{R}_N\boldsymbol{X}^{(k+1/2)}\right]$. Otherwise, use the feature sign search algorithm to solve Eq. (5), obtaining the set of sparse coefficients $\boldsymbol{\Theta}^{(k+1/2)}$.

Step 2.3: Calculate $\boldsymbol{\Theta}^{(k+1)} = T_\tau\left[\boldsymbol{\Theta}^{(k+1/2)}\right]$.

Step 2.4: Calculate $\boldsymbol{X}^{(k+1)} = \boldsymbol{\Phi} \cdot \boldsymbol{\Theta}^{(k+1)}$

Step 2.5: If mod $(k, P) = 0$, recalculate the ARM parameters for each block, and compute the similar block, update $\boldsymbol{A}$ and $\boldsymbol{B}$.

Step 2.6: Set $k = k + 1$.

Step 3: Output the SR reconstruction results $\boldsymbol{X}$.

In our algorithm, Max_Iter is used to control the number of iterations (we adopt Max_Iter =150), e is used to control the convergence degree, and P is a constant used to control the updating speed of $\boldsymbol{A}$ and $\boldsymbol{B}$. The parameter τ in Step 2.3 is selected according to the method in Reference [17], and the calculation of $\tau_{i,j}$ is based on the sparse coefficients of the similar block.

5. **Experimental Results.** In this paper, we evaluate the proposed method based on different types of LR images. For color images, we perform the our reconstruction scheme on the luminance component, while perform the bi-cubic interpolation scheme on color components Cb and Cr. In the dictionary training process, we take the 7×7 small blocks for training, and 200 classes are used for the ARM parameter model training. To obtain high frequency images, we first perform low pass filtering (the variance is 1.6) on SR images, and then subtract the Guassian filtered images from the original images. In Eq. (21), the balance factors η and γ are set to 0.04 and 0.008 respectively. In order to eliminate the blocking effect, the image blocks are divided in an overlapped manner by 4 pixels. In order to verify the effectiveness of the proposed algorithm, we compare our algorithm with the bi-cubic interpolation scheme, the ASDS reconstruction algorithm in [17] and the sparse representation based method in [16].

Fig. 1 compares four methods in terms of reconstruction visual effect. The first column shows the input LR images, the second column shows the reconstruction results by the bi-cubic interpolation scheme, the third column shows the reconstruction results by the method in [16], the fourth column shows the reconstruction results by the method in [17], and the fifth column shows the reconstruction results by our proposed method. From these results, we can see that the bi-cubic interpolation scheme obtains the most blurred results, and the algorithm in [16] cannot sufficiently reconstruct the edge information and thus the image visual effect is worse than the algorithm in [17]. For the "Butterfly" image, the algorithm in [17] has some distortion, while our algorithm can reconstruct the images with the best quality.

To objectively evaluate the four SR reconstruction methods, Table 1 shows their PSNR (Peak Signal to Noise Ratio) and SSIM (Structure SIMilarity index) results for 9 test images. From Table 1, we can see that the Bicubic algorithm achieves the lowest PSNR and SSIM values. Compared with the Bicubic algorithm, the PSNR and SSIM values of the algorithm in [16] are greatly improved, and the algorithm in [17] and our algorithm can obtain the highest PSNR and SSIM values. If we do not consider ARM and NL-M, our algorithm is generally better than the algorithm in [17] for most images. If we consider NL-M and ARM, our algorithm is slightly better than the algorithm in [17] for most images.

6. **Conclusions.** Based on the deep study of sparse representation and various image a prior information, this paper puts forward a SR image reconstruction algorithm which

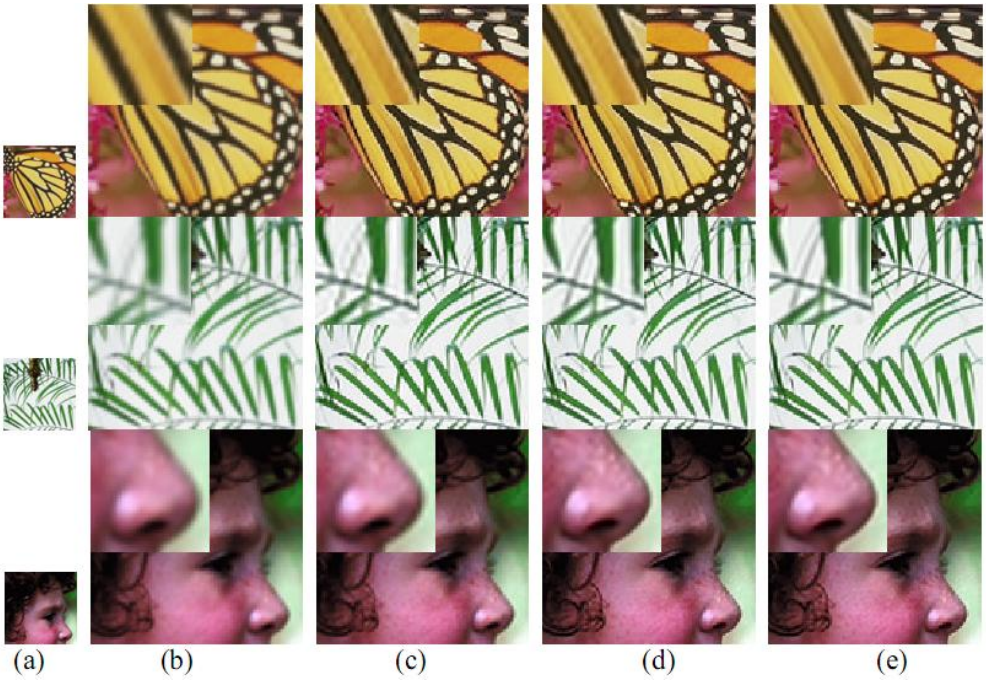

FIGURE 1. Comparisons of visual effect among various reconstruction algorithms with the reconstruction factor 3. (a) LR images; (b)Results by the Bicubic scheme; (c) Results by the method in [16]; (d) Results by the method in [17]; (e) Results by the proposed scheme.

uses the redundant dictionaries for sparse representation and adopt the iterative shrinkage approach to solve the reconstruction equation. First, the redundant dictionary and encoding dictionary are trained based on the image database. During the reconstruction process for LR images, we use the redundant dictionary to perform sparse representation on image blocks, where the sparse coefficients are respectively solved by using the coding dictionary method and the feature sign search method, and the reconstruction equation is finally solved by using the iterative shrinkage approach, obtaining the final target SR image. Simulation results show that, we can achieve better effects based on the redundant dictionary based sparse representation method. To further improve the reconstruction results, we add the autoregressive model and the non-local mean filter as additional constraints to the reconstruction equation, which can also obtain better results than the PCA sub dictionary based method.

Acknowledgement. This work was partially supported by the Zhejiang Provincial Natural Science Foundation of China under grant R1110006 and Natural Scientific Research Innovation Foundation in Harbin Institute of Technology under grant HIT. NSRIF. 2011115. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, Joint map registration and high-resolution image estimation using a sequence of undersampled images, *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1621-1633, 1997.

[2] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, Fast and robust multiframe super-resolution, *IEEE Trans. Image Processing*, vol. 13, no. 10, pp. 1327-1344, 2004.

[3] M. Protter, M. Elad, H. Takeda, and P. Milanfar, Generalizing the non-local-means to super-resolution on reconstruction, *IEEE Trans. Image Processing*, vol. 18, no. 1, pp. 36-51, 2009.

TABLE 1. Comparisons of PSNR and SSIM among various reconstruction algorithms with the reconstruction factor 3.

Image	Indicator	The Bicubic algorithm	The algorithm in [16]	The algorithm in [17] without considering ARM and NL-M	The proposed algorithm without considering ARM and NL-M	The algorithm in [17] considering ARM and NL-M	The proposed algorithm considering ARM and NL-M
Girl	PSNR	29.81	33.29	33.40	33.49	33.47	33.54
	SSIM	0.7312	0.8108	0.8191	0.8239	0.8200	0.8241
Butterfly	PSNR	20.72	26.05	26.49	26.55	27.30	27.21
	SSIM	0.7220	0.8793	0.8823	0.8882	0.9049	0.9055
Leaves	PSNR	19.81	25.40	26.09	26.21	26.80	26.81
	SSIM	0.6433	0.8713	0.8903	0.8917	0.9058	0.9075
Bike	PSNR	20.71	23.90	24.31	24.30	24.60	24.55
	SSIM	0.5725	0.7645	0.7844	0.7853	0.7959	0.7916
Flower	PSNR	24.73	28.50	28.89	28.94	29.17	29.19
	SSIM	0.6717	0.8304	0.8380	0.8419	0.8463	0.8492
Hat	PSNR	27.13	30.49	30.82	30.78	30.92	30.94
	SSIM	0.7832	0.8630	0.8681	0.8695	0.8713	0.8716
Parrots	PSNR	25.27	29.19	29.71	29.60	30.09	29.81
	SSIM	0.8164	0.9001	0.9065	0.9044	0.9101	0.9081
Plants	PSNR	27.76	32.47	33.08	33.22	33.41	33.48
	SSIM	0.7843	0.8988	0.9022	0.9081	0.9072	0.9126
Raccoon	PSNR	26.35	28.85	29.11	29.08	29.23	29.17
	SSIM	0.6285	0.7574	0.7596	0.7580	0.7655	0.7632

[4] H. S. Hou, and H. C. Andrews, Cubic spline for image interpolation and digital filtering, *IEEE Trans. Signal Processing*, vol. 26, no. 6, pp. 508-517, 1978.

[5] W. T. Freeman, T. R. Jones, and E. C. Pasztor, Example-based super-resolution, *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56-65, 2002.

[6] K. Zhang, X. Gao, X. Li, and D. Tao, Partially supervised neighbor embedding for example-based image super-resolution, *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 2, pp. 230-239, 2010.

[7] J. Yang, J. Wright, T. S. Huang, and Y. Ma, Image super resolution as sparse representation of raw image patches, *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.

[8] S. P. Belekos, N. P. Galatsanos, A. K. Katsaggelos, Maximum a posteriori video super-resolution using a new multichannel image prior, *IEEE Trans. Image Processing*, vol. 19, no. 6, pp. 1451-1464, 2010.

[9] T. Ogawa, and M. Haseyama, Missing intensity interpolation using a kernel PCA-based POCS algorithm and its applications, *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 417-432, 2011.

[10] H. Song, X. He, W. Chen, and Y. Sun, An improved iterative back-projection algorithm for video super-resolution reconstruction, *Proc. of the Symposium on Photonics and Optoelectronic(SOPO)*, pp. 1-4, 2010.

[11] J. Sun, J. Sun, Z. Xu, and H. Y. Shum, Gradient profile prior and its applications in image super-resolution and enhancement, *IEEE Trans. Image Processing*, vol. 20, no. 6, pp. 1529-1542, 2011.

[12] J. Sun, N. N. Zheng, H. Tao, and H. Y. Shum, Image hallucination with primal sketch priors, *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 729-736, 2003.

[13] M. Elad, M. A. T. Figueiredo, and Y. Ma, On the role of sparse and redundant representations in image processing, *Proceedings of IEEE*, vol. 98, no. 6, pp. 972-982, 2010.

[14] M. Elad, and M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, *IEEE Trans. on Image Processing*, vol. 15, no. 12, pp. 3736-3745, 2006.

[15] J. Mairal, G. Sapiro, and M. Elad, Learning multiscale sparse representations for image and video restoration, *Multiscale Model & Simulation*, vol. 1, no. 1, pp. 214-241, 2008.

[16] J. Yang, J. Wright, T. S. Huang, and Y. Ma, Image super-resolution via sparse representation, *IEEE Trans. image processing*, vol. 19, no. 11, pp. 2861-2873, 2010.

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

[18] H. Lee, A. Battle, R. Raina, and A. Y. Ng, Efficient sparse coding algorithms, *Advances in Neural Information Processing Systems*, pp. 801-808, 2007.

[19] C. Basso, M. Santoro, A. Verri, and S. Villa, PADDLE: proximal algorithm for dual dictionaries learning, *Proc. of the 21th International Conference on Artificial Neural Networks*, vol. 6791, pp. 379-386, 2011.

[20] B. Olshausen, and D. Field, Sparse coding with an overcomplete basis set: A strategy employed by V1 ,*Vision Research*, vol. 37, pp. 3311-3325, 1997.

[21] J. A. Tropp, and S. J. Wright, Computational methods for sparse solution of linear inverse problems, *Proceedings of IEEE*, vol. 98, no. 6, pp. 948-958, 2010.