

# Bidirectionally aligned sparse representation for single image super-resolution

Chao Xie<sup>1,2</sup> · Weili Zeng<sup>3</sup> · Shengqin Jiang<sup>1,2</sup> ·  
Xiaobo Lu<sup>1,2</sup>

Received: 30 September 2016 / Revised: 16 February 2017 / Accepted: 4 April 2017 /  
Published online: 19 April 2017  
© Springer Science+Business Media New York 2017

**Abstract** It has been demonstrated that the sparse representation based framework is one of the most popular and promising ways to handle the single image super-resolution (SISR) issue. However, due to the complexity of image degradation and inevitable existence of noise, the coding coefficients produced by imposing sparse prior only are not precise enough for faithful reconstructions. In order to overcome it, we present an improved SISR reconstruction method based on the proposed bidirectionally aligned sparse representation (BASR) model. In our model, the bidirectional similarities are first modeled and constructed to form a complementary pair of regularization terms. The raw sparse coefficients are additionally aligned to this pair of standards to restrain sparse coding noise and therefore result in better recoveries. On the basis of fast iterative shrinkage-thresholding algorithm, a well-designed mathematic implementation is introduced for solving the proposed BASR model efficiently. Thorough experimental results indicate that the proposed method performs effectively and efficiently, and outperforms many recently published baselines in terms of both objective evaluation and visual fidelity.

**Keywords** Single image super-resolution · Sparse representation · Sparse coefficient alignment · Bidirectional similarities

---

✉ Xiaobo Lu  
xblu2013@126.com

<sup>1</sup> School of Automation, Southeast University, Nanjing 210096, China

<sup>2</sup> Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Nanjing 210096, China

<sup>3</sup> College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

## 1 Introduction

Image super-resolution (SR) is by far a flourishing branch of image processing, concerning the particular issue of image resolution enhancement. When given one or more low-resolution (LR) observation(s) of the same scene, SR is capable of reconstructing a visually pleasing high-resolution (HR) output (i.e., containing more image details) [30], which might be a crucial preprocessing procedure in a wide range of digital imaging applications, such as medical diagnosis, remote sensing, and vehicle license plate recognition [53], to name but a few. Thus this technique provides us with a much economical and promising way to transcend the inherent limitations of LR optical imaging systems in place of utilizing sensor manufacturing technology, making it one of the most appealing research areas for image processing experts.

Numerous SR algorithms have been proposed over the last three decades or so. Judging from the domain employed, SR can be directly divided into two families: frequency domain approaches and spatial domain approaches. The initial research on SR [5, 24, 25, 42] belongs to the former. Although frequency domain approaches are theoretically and computationally simple, they can only be applied to pure translational model and are extremely susceptible to model errors according to the definition, which severely limits their prevalence [33]. For this reason, more and more researchers started to reconsider SR issue in spatial domain. Similarly, the proposed diverse spatial domain approaches can be roughly classified into two categories again in terms of the number of LR inputs required, i.e., multiple images SR and single image SR. For the multiple inputs case, a variety of approaches have been developed from different points of view, such as iteration back projection (IBP) based methods [18, 36], projection onto convex sets (POCS) based methods [1, 37], maximum a posteriori (MAP) estimation based methods [23, 43], regularization based methods [20, 26, 31, 32, 52, 54, 58], etc. However, it has been pointed out that the performance of this kind of methods all degrades dramatically under three circumstances where (a) the amount of LR inputs is inadequate; (b) the estimate of motion is imprecise; or (c) the scale factor increases [2, 21, 27, 45, 46].

The aforementioned limitations can be broken through by the way of exploring the other type of SR issue, i.e., single image SR (SISR). Clearly, SISR is the extreme case of SR when there is only one LR observation. Due to the ultra-insufficiency of input information, most of the proposed methods are built on the basic concept of Freeman et al. [22] assuming that the high frequency details lost in a LR image can be predicted or hallucinated by learning the co-occurrence relationship between LR patches and their corresponding HR patches extracted from a training image set, so they can be called the example or learning based methods. One group of them is neighbor embedding (NE) based methods [10, 29] which were first explored by Chang et al. These methods are based on the assumption in machine learning that small patches in LR and HR images form two manifolds lying in two distinct feature spaces but with similar local geometric structures. Thus, the SR output can be estimated by weighted summation of the  $K$  nearest HR neighbors found in the corresponding HR training database. However, the heavy running time of this method was ignored despite the fact that it is very important to real applications. Timofte et al. [39–41] incorporated the concept of learning and calculating a set of sparse dictionaries [55] (projection matrices) beforehand into the basic framework of NE to accelerate its computing process.

Another group of learning based methods, which is most relevant to this paper, is derived from the theory of sparse representation that most natural images are sparse or compressible actually when represented in the proper basis [7]. The representative work of sparse representation based SISR was proposed by Yang et al. [49, 50]. In their papers, a universal pair of HR dictionary and its corresponding LR dictionary is learned first by extracting raw patches randomly from some training images, and then sparse coding process is applied to the overlapping patches sampled in the input LR image with a raster-scan order to get the sparse coefficients. Finally, the SR output is recovered by averaging all the overlapping HR patches produced by the product of HR dictionary and coefficients. This scheme is proved to lead to a state-of-the-art result at that time, whereas it is very time-consuming to obtain two large dictionaries by randomly sampling. Hence, Zeyde et al. [55] put forward an improved method in which dimensionality reduction method is applied to the raw patches first to accelerate the subsequent dictionary learning process. To employ the priors of the training patches, clustering techniques were also introduced into the framework of sparse representation based SISR. For instance, Yang et al. [51] and Dong et al. [12] both utilize *K*-means algorithm to cluster the training raw patches into dozens of groups from which the multiple dictionaries can be learned. The superiority of their multiple dictionaries to the universal dictionary is experimentally validated.

Even though dictionary learning is an important procedure of sparse representation based SISR and the aforementioned progress actually has been made by studying on it, some more recent work [13, 14, 28, 34, 35, 47, 48] indicates that the accuracy of the sparse coefficients produced in the sparse coding process are more helpful to the performance of SISR. However, due to the complexity of the model of image degeneration, it still remains a challenging work to recover the ideal coefficients as precise as possible. Several pioneering studies on this aspect have already been made, for example, Peleg et al. [34] suggested utilizing a statistical prediction model in which a more accurate set of HR coefficients is predicted from their corresponding LR ones via the minimum mean-square error estimator. Moreover, Dong et al. [14] proposed a nonlocally centralized sparse representation (NCSR) model where the calculated coefficients are additionally centralized to a set of good estimates obtained by exploiting the nonlocal similarity within the observed image. By doing so, the two models both get improved greatly and the latter has even provided the leading SISR performance to date.

However, in light of the discovery [57] that similarities exist not only among columns but also among rows if a cluster of similar image patches is arranged in matrix form, we believe that the capacity of NCSR model is limited as it considers the column similarity only while ignores that among rows. Thus, in this paper we propose an enhanced SISR model based on bidirectionally aligned sparse representation (BASR). In our model, a pair of regularization terms is created first by exploiting both the column and row similarities (i.e., bidirectional similarities). Then, after sparse coding process, the roughly calculated sparse coefficients are simultaneously aligned to the pair of terms in order to compensate the errors caused by image noise and degradation, and consequently increase the accuracy of the sparse coefficients and SISR performance. Furthermore, for a more rapid convergence, the fast iterative shrinkage-thresholding algorithm (FISTA) [3] is adopted in this paper instead of employing the conventional iterative thresholding algorithm (ITA) [11]. Extensive experiments have demonstrated

that our proposed BASR model outperforms its recent counterparts in terms of both visual quality and numerical evaluations.

The rest of this paper is organized as follows: In section 2, we formalize the sparse representation based SISR problem. Section 3 presents the proposed BASR model for SISR issue and its implementation in detail. Experimental results and analysis are given in Section 4, while Section 5 concludes the whole paper.

## 2 Problem formulation

The goal of SISR can be regarded as recovering the potential HR image as precise as possible from just only one LR input. For a comprehensive analysis, the first step is to set up a suitable single image degradation model which relates the original HR image to the observed and degraded LR image.

Assume that  $\mathbf{X}$  is an ideal HR image, while  $\mathbf{Y}$  is the corresponding LR image of  $\mathbf{X}$  in the same scenario. Then both of them are lexicographically rearranged into vector form, i.e.,  $\mathbf{X} \in R^N$ ,  $\mathbf{Y} \in R^M$ , where  $N > M$ ,  $r^2 = N/M$ , and  $r$  is the scale factor. The degradation can be typically described as [38]

$$\mathbf{Y} = \mathbf{SBX} + \mathbf{V} = \mathbf{HX} + \mathbf{V} \quad (1)$$

where  $\mathbf{S} : R^N \rightarrow R^M$  is the down-sampling operator,  $\mathbf{B} : R^N \rightarrow R^N$  is the blurring operator,  $\mathbf{V} \in R^M$  is the additive noise, and  $\mathbf{H} : R^N \rightarrow R^M$  is the degradation operator which can be viewed as a composite operator of both  $\mathbf{S}$  and  $\mathbf{B}$ .

Clearly, the fundamental constraint of SISR is that the recovered image should approximately reproduce the LR observation after imposing the same degradation on it. Nevertheless, since too much information is discarded during the high-to-low acquisition process, the linear eq. (1) is seriously underdetermined, i.e., infinitely many solutions may be suitable for (1). In order to obtain an appropriate solution, the researchers in [17, 49, 50] set up the initial framework of sparse representation based SISR, which incorporates the sparsity prior and the local-to-global reconstruction concept to the basic constraint. To be specific, suppose that the operator  $\mathbf{R}_i : R^N \rightarrow R^m$  is used to extract the  $i$ -th patch of size  $\sqrt{n} \times \sqrt{n}$  from an  $N$  length image and vectorize it, thus the  $i$ -th patch of  $\mathbf{X}$  can be readily expressed as  $\mathbf{x}_i = \mathbf{R}_i \mathbf{X}$ . With the corresponding observed LR patch  $\mathbf{y}_i$ , each patch  $\mathbf{x}_i$  can be sparsely represented by the formula  $\mathbf{x}_i = \mathbf{D} \alpha_{y,i}$ , where  $\alpha_{y,i}$  is its sparse coefficient (representation) that can be calculated by a sparse coding operation with respect to a known and proper dictionary  $\mathbf{D}$

$$\alpha_{y,i} = \arg \min_{\alpha_i} \left\{ \|\mathbf{y}_i - \mathbf{H} \mathbf{D} \alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \right\} \quad (2)$$

where  $\lambda$  is a trade-off to make a balance between the two terms. Note that the  $l_1$ -norm term is a regularization term representing the sparsity prior and it has already been changed from  $l_0$ -norm to  $l_1$ -norm as long as the coefficients are sufficiently sparse due to non-convex character of  $l_0$ -norm [8, 9, 16].

Usually, we work with overlapped patches to suppress the boundary artifacts. Under the circumstances of maximal overlaps, a total number of  $Q = (\sqrt{N} - \sqrt{n} + 1)^2$  patches can be represented. After imposing a global constraint on these patches, the optimal reconstruction of the whole HR image  $X$  can be straightforwardly computed by averaging all the obtained local patches according to [17]

$$X = D \circ \alpha_Y \triangleq \left( \sum_{i=1}^Q R_i^T R_i \right)^{-1} \left( \sum_{i=1}^Q R_i^T D \alpha_{y,i} \right) \quad (3)$$

where  $\alpha_Y$  represents the concatenation of all sparse coefficients, and a shorthand notation “ $\circ$ ” is defined here for a briefer expression in the following parts.

Eqs. (2) and (3) can be reformulated together into a more unified formation to stand for the sparse coding process and local-to-global reconstruction simultaneously

$$\alpha_Y = \arg \min_{\alpha} \left\{ \|Y - HD \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 \right\} \quad (4)$$

where the first term corresponds to the data fidelity constraint, and the second one corresponds to the sparsity prior constraint.

In summary, under the basic framework of sparse representation, the ill-posed SISR problem is further regularized by the sparse prior of patches in addition to the local and global data fidelity constraint, resulting in a proper and stable solution.

### 3 Proposed BASR model for SISR

In this section, we are going to present an enhanced model, namely BASR, which is designed for handling the SISR issue under the particular circumstances where no external databases is allowed for prior or dictionary learning. The presentation of BASR model begins with the modeling of bidirectional similarities that is the theoretical support of subsequent content. Once it is finished, the key process, sparse coefficient alignment, is able to be established in both directions. Afterward, the way of dictionary learning and detailed implementation are specified sequentially.

#### 3.1 Modeling of bidirectional similarities

In this paper, the proposed bidirectional similarities consist of the row similarity and column similarity. To construct the bidirectional similarities, the first step is to establish the similarity data matrix [57]. For each patch  $x_i$ , we search for its  $P$  closest counterparts (include itself) from the whole HR image  $X$ , in the sense of Euclidean distance metric. By concatenating the patch and its counterparts together, we can obtain a matrix  $S_i \in R^{n \times P}$ , namely the similarity data matrix of the  $i$ -th patch. As we mentioned before, patch similarity can be found not only among the columns of similarity data matrix but also among the rows of it. Therefore, the next step is to exploit the column similarity and row similarity, respectively, by virtue of the similarity data matrix.

As for column similarity, it was first put forward in the non-local means (NLM) algorithm [6] and applied to the application of image denoising. But, different from the original NLM algorithm that actually means a weighted average of all similar patches; we determine to use every column of  $\mathcal{S}_i$  as dictionary atoms to approximately represent the corresponding patch. This process can be formulated as

$$\beta_i = \arg \min_{\beta_i} \left\{ \|\mathbf{x}_i - \mathcal{S}_i \beta_i\|_2^2 + \eta_1 \|\beta_i\|_2^2 \right\} \quad (5)$$

where  $\beta_i$  is called the column similarity coefficient of the  $i$ -th patch, and  $\eta_1$  is its regularization parameter.

In (5),  $l_2$ -norm of the coefficient is designed as the regularization term, which is involved for the purpose of alleviating the singularity in calculation and avoiding the trivial solution. As you may see, the formula is actually in the same form of classic Tikhonov regularization, which is also known as ridge regression in statistics. Its explicit solution can be easily given by

$$\beta_i = (\mathcal{S}_i^T \mathcal{S}_i + \eta_1 \mathbf{I})^{-1} \mathcal{S}_i^T \mathbf{x}_i \quad (6)$$

where  $\mathbf{I}$  represents the identical matrix.

When turning to exploit row similarity, we are motivated by the concept of piecewise autoregressive (AR) model. As a classic but powerful method in statistics, AR model has been successfully employed in some other image processing applications, such as image interpolation [56] and image denoising [57]. The key point in it is that if a natural image is tailored into small local parts, each part can be viewed as one stationary process. In other words, it suggests that natural images are piecewise stationary and able to be modeled by a set of AR models.

Therefore, in this work we assume that the central pixel of each patch can be linearly represented by its neighboring pixels with the coefficient calculated by an AR model. Moreover, the patches that belong to the same similarity data matrix should have identical AR coefficient as they share the same similarity. Let  $\mathbf{C}$  be the operator to extract the central pixels from the similarity data matrix (i.e., the central row of it). Similar to column similarity, this progress can also be modeled as

$$\gamma_i = \arg \min_{\gamma_i} \left\{ \|(\mathbf{C}\mathcal{S}_i)^T - \mathcal{S}_i^T \gamma_i\|_2^2 + \eta_2 \|\gamma_i\|_2^2 \right\} \quad (7)$$

where  $\alpha_i$  is called the  $i$ -th row similarity coefficient, and the closed-form solution is given by

$$\gamma_i = (\mathcal{S}_i \mathcal{S}_i^T + \eta_2 \mathbf{I})^{-1} \mathcal{S}_i (\mathbf{C}\mathcal{S}_i)^T \quad (8)$$

With column and row coefficients, the similarities in both directions can be calculated as  $\{\mathcal{S}_i \beta_i\}_{i=1}^Q$  and  $\{\mathcal{S}_i^T \gamma_i\}_{i=1}^Q$ , respectively. Be aware that each calculated row similarity is composed of the central pixels of the corresponding similar patches, so it is incorrectly ordered with what we have requested. Therefore, we need to spread it out in the whole HR image first,

and then rearrange it in the same order as  $\alpha$ . After doing so, the rearranged column and row similarities are denoted as  $\varphi_i'$  and  $\psi_i'$ , respectively, which are the proposed bidirectional similarities in this paper. By taking advantage of the bidirectional similarities, it allows us to structure a much more accurate sparse representation model, which will be introduced specifically in the following subsection.

### 3.2 Bidirectional sparse coefficient alignment

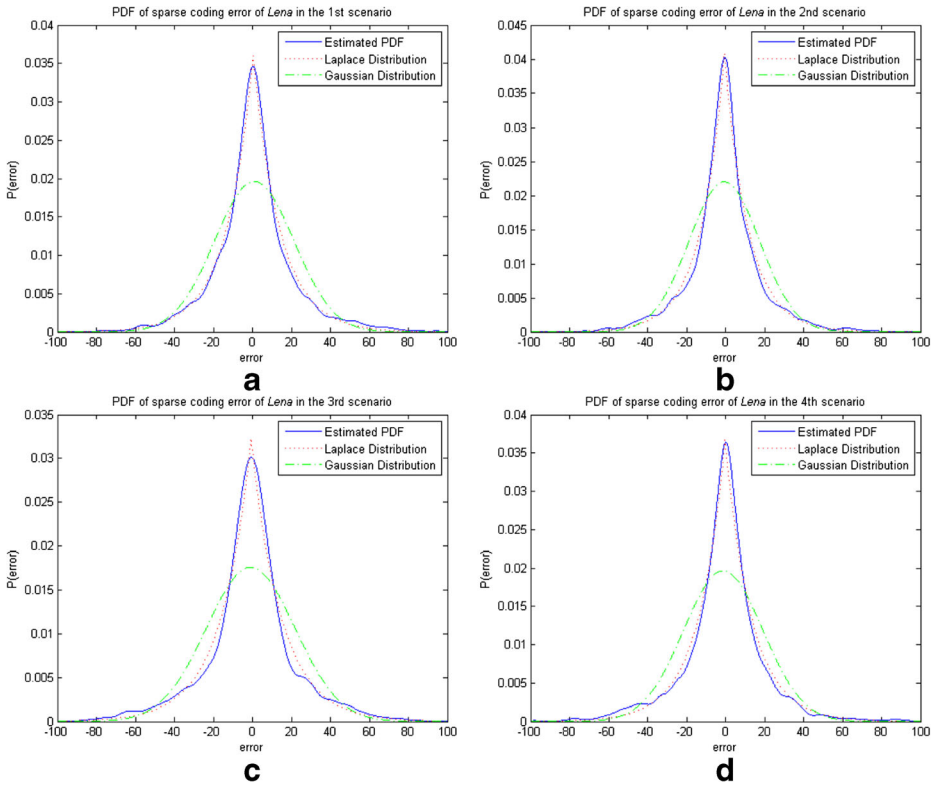
As mentioned before, it has been found that the accuracy of sparse coefficients is of great significance to sparse representation based SISR. Nevertheless, model (4) which uses the sparse prior only may not lead to a precise enough output due to the complexity of image degradation. And it can be expected that a performance enhancement would be acquired by suppressing the sparse coding error caused by degradation and noise. Thus, in this subsection we propose an enhanced model, in which the roughly computed sparse coefficients are simultaneously aligned to the bidirectional similarities proposed previously.

Before the construction of our model, an important procedure we need to do is to change the values of bidirectional similarities from pixel domain to sparse coefficient domain through a sparse coding process. As will be introduced in the next subsection, all the sub-dictionaries adopted in this paper are orthogonal, so the coding process is simplified to only multiplying the pixel values by the transpose of the corresponding sub-dictionary. Given the sub-dictionary  $D_i$  of the  $i$ -th patch, this process is formulated as  $\varphi_i = D_i^T \varphi_i'$  and  $\psi_i = D_i^T \psi_i'$ . By incorporating this pair of similarities into the basic model (4) as additional regularization terms, we get the objective function of our proposed BASR model

$$\alpha = \arg \min_{\alpha} \left\{ \|Y - HD \circ \alpha\|_2^2 + \lambda \|\alpha\|_1 + \mu_1 \|\alpha - \varphi\|_p^p + \mu_2 \|\alpha - \psi\|_p^p \right\} \quad (9)$$

where  $\varphi$  and  $\psi$  are the concatenation of all  $\varphi_i$  and  $\psi_i$  respectively, representing the bidirectional similarities of  $\alpha$ , respectively.

As you can see, in our BASR model, the output sparse coefficients are not only of the characteristic of sparsity, but also bidirectionally aligned so that the errors caused by degradation and noise can be efficiently suppressed. Furthermore, similar to [13], a more comprehensive analysis of sparse coding error is conducted and provided here in order to illustrate its statistical property, and consequently determine the type of norm to be used in (9). Specially, the test images *Lena* is chosen as an HR sample from which its four degraded versions are able to generate by applying the degradations specified in subsection 4.1. That is to say, here, we take into consideration all the four SISR scenarios. Using the given sub-dictionaries, it is straightforward to calculate the difference between the ideal and estimated sparse coefficients, namely the sparse coding error, by solving (4). Be aware that, to be fully convincing, the parameters used here are set to be identical with what will be adopted in the experimental section. Eventually, the probability density functions (PDFs) of sparse coding error under the four scenarios are plotted in Fig. 1(a)–(d) with respect to the 5th, 10th, 15th, 20th sub-dictionary, respectively. As shown in Fig. 1, the estimated PDFs are unable to conform to the Gaussian distributions,



**Fig. 1** PDFs of sparse coding errors of image *Lena* in (a) scenario 1, (b) scenario 2, (c) scenario 3 and (d) scenario 4

but they all can fit in well with the Laplace distributions. Therefore,  $l_1$ -norm should be picked to model the sparse coding error (i.e.,  $p = 1$ ), motivated by the analysis conducted.

### 3.3 Dictionary learning and adaptive selection

Clearly, two key procedures left undone are dictionary learning and adaptive selection of one dictionary for each local patch. Conventional way of dictionary learning aims at learning a universal and over-complete dictionary to code different varieties of local structures [49, 50, 55]. However, recently it has been proved that sparse coding process under this kind of dictionaries is inherently time-consuming and potentially unstable [19]. Thus we turn to the help of another promising strategy, namely adaptive sparse domain selection (ASDS) [12].

Originally, ASDS needs an extra database of raw image patches to train on, whereas in this paper we are considering a more practical situation where no external information is available. To overcome this, an alternative training database is constructed for ASDS



by sampling patches from the currently estimated HR image and its down-scaled versions instead. With training database, the specific produces goes as follows: firstly, the training patches are gathered and partitioned into  $K$  clusters via  $K$ -means clustering. After applying PCA to each cluster, we can totally obtain  $K$  orthogonal and compact sub-dictionaries which compose the final dictionary of this paper, denoted by  $\{\mathbf{D}_i\}_{i=1}^K$ . Then, for an input patch  $\mathbf{x}_i$  to be coded, the sub-dictionary  $\mathbf{D}_i$  belonging to the nearest cluster is selected from the overall dictionary, and the sparse coding process is greatly simplified to matrix multiplication of the form:  $\alpha_i = \mathbf{D}_i^T \mathbf{x}_i$ . Since each given patch can be better represented by the adaptively selected sub-dictionary, the whole recovered image is more accurate than just using a universal dictionary. Moreover, this learning and coding strategy implicitly enforces the coefficient of the given patch with respect to the other sub-dictionaries equal to zero. That is to say, our model guarantees the local sparsity of coefficients spontaneously, thus the regularization term in (9) enforcing local sparsity can be omitted. The objective function finally becomes

$$\alpha = \arg \min_{\alpha} \left\{ \|\mathbf{Y} - \mathbf{H}\mathbf{D} \circ \alpha\|_2^2 + \mu_1 \|\alpha - \varphi\|_1 + \mu_2 \|\alpha - \psi\|_1 \right\} \quad (10)$$

### 3.4 Summary and mathematic implementation

It can be seen that the proposed BASR model (10) is a hybrid optimization problem with the co-occurrence of  $l_2$ -norm and  $l_1$ -norm, which makes it become non-convex and have no closed-form solution. Therefore, in our implementation, the proposed model is designed to be iteratively solved in a patchwise manner. Without loss of generality, this model can be rewritten into a patchwise form

$$\{\alpha_i\}_{i=1}^Q = \arg \min_{\alpha_i} \left\{ \|\mathbf{y}_i - \mathbf{H}\mathbf{D}_i \alpha_i\|_2^2 + \mu_1 \|\alpha_i - \varphi_i\|_1 + \mu_2 \|\alpha_i - \psi_i\|_1 \right\}, \text{for } i = 1, 2, \dots, Q \quad (11)$$

On the basis of fast iterative shrinkage-thresholding algorithm (FISTA) [3], a local-to-global and coarse-to-fine solving course is able to be concisely expressed as follows: (for more details about the whole process, please refer to Appendix A).

$$\begin{aligned} \left\{ \alpha_i^{(l+1)} \right\}_{i=1}^Q &= \rho(\alpha_i^{\text{tmp}}, L) = \rho(\mathbf{D}_i^T \mathbf{R}_i \mathbf{X}^{\text{tmp}}, L) = \rho(\mathbf{D}_i^T \mathbf{R}_i \cdot h(\mathbf{X}^{(l)}, \mathbf{X}^{(l-1)}), L), \text{for } i \\ &= 1, 2, \dots, Q \end{aligned} \quad (12)$$

Notation  $L$  is a scalar involved here to control the magnitude of step-size and its value can be determined by employing a backtracking step-size rule, function  $\rho$  is cast as the shrinkage operator which is defined in Appendix A. Moreover, to accelerate the iteration, the temporary variable before shrinkage operation in (12) ought not to be computed by considering the result obtained in the previous iteration only (i.e.,  $\mathbf{X}^{(l)}$ ), but rather calculated by utilizing a very special linear combination of the previous two results (i.e.,  $\mathbf{X}^{(l)}$  and  $\mathbf{X}^{(l-1)}$ ), which is conveyed by function  $h$ . In summary, the detailed

description of the implementation in this paper is outlined in Algorithm 1, while the block diagram of proposed BASR method is illustrated in Fig. 2.

---

**Algorithm 1:** Mathematic implementation of the BASR model for SISR

---

1. **Input:** LR image  $Y$  and scale factor  $r$ .

2. **Initialization:**

- a) The initial estimate of  $X$ , denoted by  $\hat{X}^{(0)}$ , is calculated by applying bicubic interpolation to the input LR image  $Y$ . Set  $\hat{X}^{(1)} = \hat{X}^{(0)}$ .
  - b) Set the parameters of BASR:  $\mu_1, \mu_2, \eta_1$  and  $\eta_2$ . Set the initial values of FISTA:  $t^{(0)} = t^{(1)} = 1$ , the maximum outer and inner iteration numbers  $I_1$  and  $I_2$ , and reset iteration counter  $l = 0$ . Set the initial values of backtracking:  $\delta = 1.5$  and  $L = 1$ .
- 

3. **Iteration:** outer loop (dictionary learning and similarity modeling) iterates from 1 to  $I_1$ .

- a) Construct the database of patches with respect to  $\hat{X}^{(l)}$ , and train the dictionary  $\{D_i\}_{i=1}^K$  on this database using  $K$ -means clustering and PCA algorithms.
- b) For each local patch  $x_i^{(l)} = R_i \hat{X}^{(l)}$ , the nearest sub-dictionary  $D_{q_i}$  is selected and the similarity data matrix  $S_i$  is constructed.
- c) Inner loop (sparse coding and coefficient alignment) iterates from 1 to  $I_2$ .

- i)  $Z = \hat{X}^{(l)} + \frac{t^{(l-1)} - 1}{t^{(l)}} (X^{(l)} - X^{(l-1)})$ , then calculate  $\alpha^{(l)}$  with respect to  $Z$ .

- ii) Repeat the following steps (backtracking) while the judgment condition in s5) is not satisfied.

- s1)  $X^{imp} = Z - \frac{2}{L} (H^T H Z - H^T Y)$ .

- s2)  $\alpha^{imp} = [D_{q_1}^T R_1 X^{imp}, \dots, D_{q_{i_0}}^T R_{i_0} X^{imp}]$ .

- s3) Generate  $\varphi$  and  $\psi$  using (5) - (8).

- s4) Compute  $\alpha^{imp} = \rho(\alpha^{imp}, L)$  using (19) - (20).

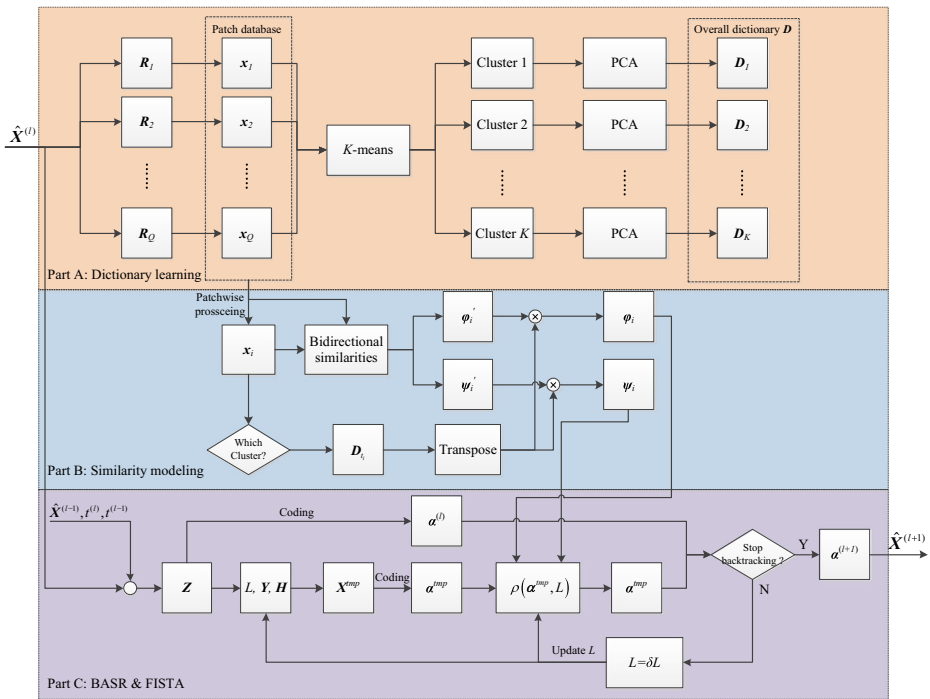
- s5) If  $f(\alpha^{imp}) + g(\alpha^{imp}) \leq \omega(\alpha^{imp}, \alpha^{(l)}, L)$  is not met, set  $L = \delta L$  and continue; otherwise break

- iii) Upgrade the whole image  $\hat{X}^{(l+1)}$  with respect to the current  $\alpha^{imp}$  using (3).

- iv)  $t^{(l+1)} = \frac{1 + \sqrt{1 + 4(t^{(l)})^2}}{2}$ .

4. **Output:** the recovered HR image  $\hat{X}$ .

As we can see in Algorithm 1, a nested iteration consisting of inner and outer loops is employed. The reason is that the main computational burden exists in the procedures of dictionary learning and similarity modeling, and the variables involved in these parts do not change drastically while iteration proceeds. Therefore, these procedures can be put in the outer loop and only need to be executed every  $I_2$  iterations in order to save computation cost. In addition, the step (i) and (iv) in the inner loop is additionally requested computation cost of FISTA. When compared to other steps, the two are both computationally negligible, thus the algorithm above remains almost the same computational burden but with a faster convergence



**Fig. 2** Block diagram of proposed BASR method. Note that part B, A, and C in the above diagram actually correspond to the specific contents of subsection 3.1, 3.3, and 3.2 + 3.4, respectively

rate. The theoretical convergence of FISTA has been well proved, please refer to [3, 11] for details.

### 4 Experimental results and analysis

In this section, a series of experiments on natural images have been designed and conducted to verify the effectiveness and robustness of our proposed BASR method in comparison with eight recent state-of-the-art counterparts, including SCSR [50], SLSR [55], SPSR [34], ASDS [12], NCSR [14], ANR [39], A+ [40] and SRCNN [15].

#### 4.1 Experimental settings

First of all, thirteen different types of genetic images (refer to Table 1) are accepted as test benchmarks, among which the first ten were presented in [12] while the rest three appeared in [34]. Then, the degradation model (1) needs to be applied to these test samples for generating the LR images that will be used as inputs in the following experiments. But be aware that two prevalent but different types of degradation configurations were used among the aforementioned eight methods: the first type tends to use the bicubic filter to blur the original image first and then down-sample it with a prespecified decimation factor in both horizontal and vertical directions; by contrast the second chose to employ a Gaussian filter instead, followed by the same down-sampling process. It can be seen that the essential difference between the two

**Table 1** Numerical results (PSNR and SSIM) of the reconstructed HR images in scenario 1

Images	SCSR [50]	SLSR [55]	SPSR [34]	ASDS [12]	NCSR [14]	ANR [39]	A+ [40]	SRCNN [15]	BASR
Bike	23.55	23.85	24.34	24.55	24.68	23.94	24.57	24.64	<b>24.89</b>
	0.7465	0.7660	0.7894	0.7980	0.8039	0.7712	0.7982	0.7968	<b>0.8109</b>
Butterfly	24.79	25.89	26.61	27.16	28.04	25.82	27.06	27.83	<b>28.75</b>
	0.8140	0.8756	0.8971	0.9026	0.9208	0.8702	0.9064	0.9084	<b>0.9291</b>
Hat	29.94	30.45	30.80	31.02	31.22	30.33	31.10	31.01	<b>31.42</b>
	0.8391	0.8600	0.8648	0.8682	0.8750	0.8579	0.8735	0.8682	<b>0.8789</b>
Leaves	24.10	25.28	25.84	26.69	27.24	25.20	26.14	26.62	<b>28.42</b>
	0.8173	0.8738	0.8928	0.9089	0.9205	0.8677	0.9032	0.9094	<b>0.9375</b>
Parrot	28.78	29.43	29.92	30.26	30.23	29.54	30.10	30.49	<b>30.85</b>
	0.8870	0.9067	0.9134	0.9153	0.9171	0.9078	0.9170	0.9179	<b>0.9229</b>
Flower	27.92	28.58	28.96	29.37	29.19	28.73	29.21	29.31	<b>29.60</b>
	0.8078	0.8336	0.8462	0.8540	0.8546	0.8381	0.8535	0.8522	<b>0.8633</b>
Girl	32.97	33.45	33.56	33.61	33.63	33.51	33.66	33.61	<b>33.79</b>
	0.8133	0.8229	0.8240	0.8256	0.8295	0.8260	0.8296	0.8294	<b>0.8326</b>
Parthenon	26.26	26.76	26.85	27.01	27.07	26.67	27.03	27.21	<b>27.30</b>
	0.7156	0.7331	0.7380	0.7434	0.7490	0.7312	0.7459	0.7502	<b>0.7555</b>
Plants	31.69	32.53	32.89	33.49	33.65	32.58	33.48	33.63	<b>34.05</b>
	0.8820	0.9027	0.9071	0.9122	0.9204	0.9046	0.9205	0.9164	<b>0.9256</b>
Raccoon	28.62	28.93	29.12	29.26	29.29	28.99	29.22	29.41	<b>29.42</b>
	0.7475	0.7583	0.7622	0.7684	0.7719	0.7618	0.7684	0.7716	<b>0.7752</b>
Lena	30.57	31.36	31.83	32.10	32.17	31.42	32.07	32.18	<b>32.36</b>
	0.8464	0.8720	0.8784	0.8840	0.8872	0.8741	0.8858	0.8862	<b>0.8901</b>
Starfish	27.34	28.04	28.70	29.13	29.08	28.07	28.52	29.09	<b>29.48</b>
	0.8268	0.8490	0.8645	0.8739	0.8749	0.8503	0.8630	0.8734	<b>0.8814</b>
Zebra	25.25	26.77	26.87	27.44	27.52	26.63	27.39	27.85	<b>28.43</b>
	0.8137	0.8583	0.8637	0.8724	0.8755	0.8565	0.8716	0.8773	<b>0.8862</b>
Average	27.83	28.56	28.95	29.31	29.46	28.57	29.20	29.45	<b>29.90</b>
	0.8121	0.8394	0.8494	0.8559	0.8616	0.8398	0.8567	0.8583	<b>0.8684</b>

The numbers in bold indicate the best performance (i.e., the highest values)

configurations is actually the choice of blur kernel. In this paper, both types of the filters should be taken into account in order to be fully convincing. But to emphasize the performance influence caused by different blur kernels, we decide to fix the decimation factor to be 3 (the benchmark value) throughout our experiments so that that part of impact introduced by down-sampling could always be the same. In addition, the additive zero-mean Gaussian noise of standard deviation 5 could be further added to the degraded inputs for testing the robustness of all methods to noise. Therefore, in total, four different SISR scenarios are taken into consideration in this paper, which are sequentially specified as follows:

Scenario 1: HR images are blurred by the Bicubic filter, followed by down-sampling.<sup>1</sup>

Scenario 2: HR images are blurred by the Gaussian filter of size  $7 \times 7$  and standard deviation 1.6, followed by down-sampling.

Scenario 3: Use the LR outputs of scenario 1 but contaminated by Gaussian noise additionally.

Scenario 4: Use the LR outputs of scenario 2 but contaminated by Gaussian noise additionally.

<sup>1</sup> This whole operation can be directly accomplished by using the MATLAB function “imresize” with the method option “bicubic”.

The other main setting of parameters involved in this paper empirically goes as follows:  $P = 12$ ,  $K = 50$ ,  $\eta_1 = 0.1$ ,  $\eta_2 = 0.3$ ,  $I_1 = 9$ ,  $I_2 = 40$  and the local patches are extracted with maximal overlaps, whose size is set to be  $5 \times 5$  (i.e.,  $n = 25$ ). Note that, according to some preliminary tests, we have found that the performance of our BASR method is almost insusceptible to the aforementioned parameters within a reasonable range, whereas the weights of regularization terms (i.e.,  $\mu_1$  and  $\mu_2$ ) play a much important role. Therefore, the determination of the regularization weights deserves a deep consideration. In order to simplify the determination, we decide to keep the weight of column similarity (i.e.,  $\mu_1$ ) equal to the value employed in [14] unvaryingly, meanwhile search for the optimal value of the other weight (i.e.,  $\mu_2$ ). In doing so, we can pay more attention on the complementary effect produced by the regularization terms of BASR model rather than get stuck in doing parameter optimization. As a result, the weights  $\mu_1$  and  $\mu_2$  are set to be 0.7 and 0.1 respectively for noiseless experiments, while under noisy condition they are equal to 1 and 0.2, respectively.

Moreover, considering that human visual system is more sensitive to the changes in luminance than to those in chromaticity, YCbCr color space is adopted for representing color images in place of RGB space. All of the competitive methods are applied to the luminance component only, while the chromaticity components are simply interpolated from the input LR image to the target HR image by bicubic interpolation. To qualitatively and quantitatively evaluate the performance, the reconstructed images produced by various methods are contrasted in terms of both visual quality and two widely used numerical indicators, i.e., PSNR and SSIM index [44]. When computing both of the numerical indicators, image borders which were neglected in [15, 38, 55] are uniformly taken into consideration in this paper for a fair comparison. All experiments were performed in the MATLAB R2013b environment on a PC with Core i5 3.2GHz CPU and 4GB RAM.

## 4.2 Experimental results

In this subsection, all competitive methods are implemented and evaluated under the four different SISR scenarios specified above. First, let us concentrate on the first two scenarios, which actually correspond to noiseless cases but with two different blur kernels. For scenario 1, the numerical results are reported in Table 1. Note that for each image its upper row in the Table shows the PSNR values (unit is dB) while the lower row provides the SSIM indexes (dimensionless quantity), and the layout is accepted in every Table of this paper. From Table 1 we can see that, in terms of the quantitative assessment, the conventional sparse representation based methods, SCSR and SLSR, give the lowest results all the time, while the other six methods are all better than them more or less. This is mainly due to that some particular progress has been made in either dictionary learning or sparse coding process. In the meantime, it is doubtless that our proposed BASR method always provides the highest numerical values among all competitive methods, since it not only inherits the good properties of sparse representation based SISR framework but also benefits quite a lot from the proposed bidirectional sparse coefficient alignment procedure. When testing environment turns to the second scenario, the whole numerical indicators are listed in Table 2, from which a consistent conclusion can be readily drawn that our BASR method is still superior to all the other counterparts, including the best competitor NCSR. To be precise, it outperforms the second best method, namely NCSR, by an average gain of 0.63 dB and 0.0086 in PSNR and SSIM index, respectively. Moreover, by comparing the numerical results of scenario 1 and 2, we noticed that the numerical differences of our method between the two scenarios ( $-0.01$  dB for

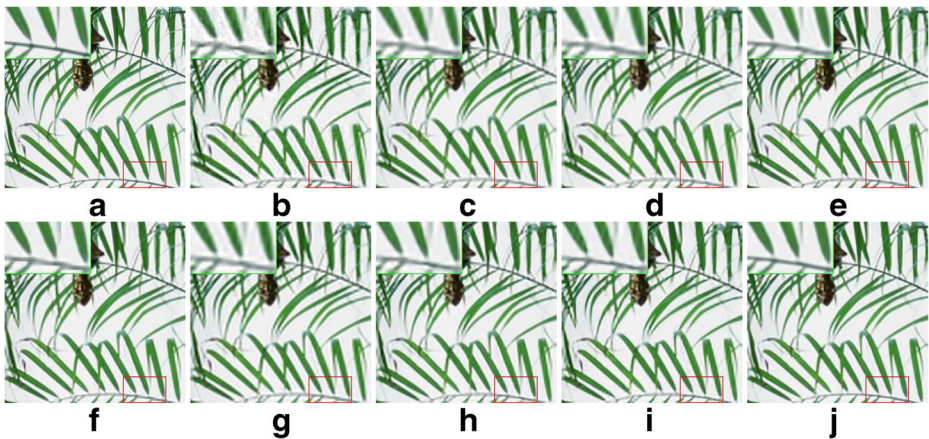
**Table 2** Numerical results (PSNR and SSIM) of the reconstructed HR images in scenario 2

Images	SCSR [50]	SLSR [55]	SPSR [34]	ASDS [12]	NCSR [14]	ANR [39]	A+ [40]	SRCNN [15]	BASR
Bike	22.79	22.72	24.22	24.58	24.60	22.74	23.04	23.10	<b>24.94</b>
	0.6920	0.6962	0.7838	0.7975	0.7997	0.6992	0.7149	0.7154	<b>0.8099</b>
Butterfly	23.90	23.99	26.53	27.18	28.02	23.86	24.37	24.46	<b>28.80</b>
	0.7983	0.8340	0.8946	0.9024	0.9145	0.8286	0.8484	0.8471	<b>0.9288</b>
Hat	29.09	29.35	30.77	30.89	31.16	29.18	29.61	29.48	<b>31.42</b>
	0.8188	0.8333	0.8633	0.8649	0.8696	0.8301	0.8400	0.8356	<b>0.8774</b>
Leaves	23.24	23.26	25.62	26.59	27.28	23.16	23.59	23.62	<b>28.35</b>
	0.7832	0.8122	0.8875	0.9065	0.9203	0.8065	0.8287	0.8272	<b>0.9366</b>
Parrot	28.00	27.92	29.90	30.37	30.32	28.00	28.30	28.32	<b>30.85</b>
	0.8689	0.8803	0.9122	0.9132	0.9132	0.8817	0.8875	0.8866	<b>0.9219</b>
Flower	27.28	27.28	28.87	29.28	29.38	27.33	27.60	27.63	<b>29.64</b>
	0.7736	0.7827	0.8437	0.8491	0.8540	0.7858	0.7961	0.7945	<b>0.8624</b>
Girl	32.34	32.49	33.47	33.53	33.62	32.54	32.64	32.66	<b>33.78</b>
	0.7870	0.7907	0.8216	0.8204	0.8269	0.7940	0.7966	0.7976	<b>0.8309</b>
Parthenon	25.85	26.01	26.84	27.00	27.13	25.96	26.18	26.23	<b>27.35</b>
	0.6840	0.6930	0.7353	0.7435	0.7494	0.6925	0.7013	0.7030	<b>0.7552</b>
Plants	30.85	31.01	32.77	33.45	33.82	30.97	31.40	31.38	<b>34.03</b>
	0.8589	0.8679	0.9048	0.9080	0.9170	0.8687	0.8789	0.8759	<b>0.9241</b>
Raccoon	28.07	28.08	28.98	29.24	29.22	28.15	28.25	28.38	<b>29.42</b>
	0.7016	0.7039	0.7564	0.7661	0.7685	0.7084	0.7125	0.7159	<b>0.7718</b>
Lena	29.74	30.04	31.69	31.99	32.12	29.99	30.35	30.31	<b>32.35</b>
	0.8248	0.8404	0.8753	0.8794	0.8821	0.8414	0.8488	0.8483	<b>0.8889</b>
Starfish	26.58	26.60	28.61	29.03	29.18	26.65	26.85	26.98	<b>29.49</b>
	0.7906	0.7985	0.8610	0.8687	0.8717	0.8017	0.8090	0.8132	<b>0.8802</b>
Zebra	24.35	24.41	26.74	27.49	27.66	24.32	24.70	24.91	<b>28.48</b>
	0.7730	0.7917	0.8594	0.8712	0.8750	0.7916	0.8018	0.8085	<b>0.8861</b>
Average	27.08	27.17	28.85	29.28	29.50	27.14	27.45	27.50	<b>29.91</b>
	0.7811	0.7942	0.8461	0.8531	0.8586	0.7946	0.8050	0.8053	<b>0.8672</b>

The numbers in bold indicate the best performance (i.e., the highest values)

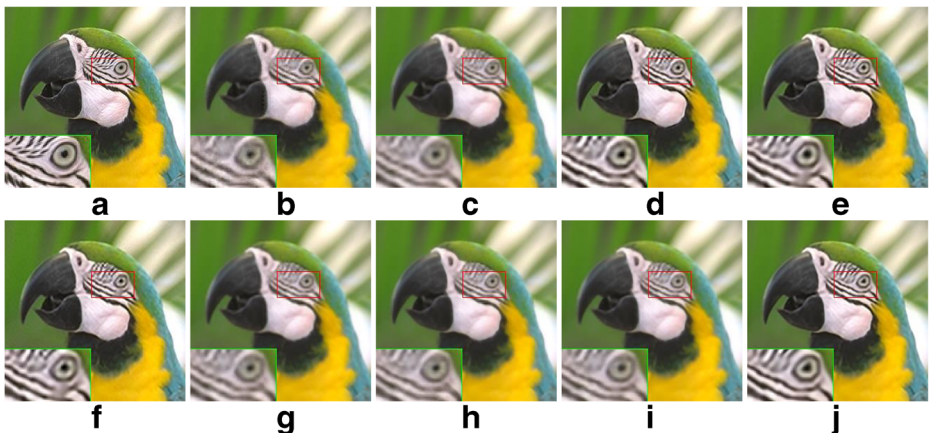
PSNR and 0.0012 for SSIM) are negligible. That is to say, our BASR model is able to provide the equal performance when the blur operator is changed from bicubic filter to Gaussian, showing a strong robustness to the variation on blur type. Next, to further evaluate the visual quality obtained by all methods, two sets of visual comparisons on image *Leaves* and *Parrot* under the first two scenarios are illustrated in Figs. 3 and 4, respectively. As shown in the figures, the proposed BASR method can produce the most visually pleasing outputs in terms of preserving more fine details and sharp edges. For instance, the stems and leaves in Fig. 3(j) are most consistent with those belonging to the original image *Leaves* showed in Fig. 3(a), while others are blurred, twisted or distorted to some degree. Meanwhile, the stripes around the eye of image *Parrot* in Fig. 4(j) are the clearest and sharpest in comparison with the other methods.

Then, in consideration of the fact that in practice the LR inputs of SISR are often contaminated by noise, thus we move on to the next two SISR scenarios for the purpose of testing the robustness of these methods to noise. Gaussian white noise is further added to the same LR images previously used in scenarios 1 and 2 to generate the new inputs for scenarios 3 and 4 respectively, making the task more challenging. Tables 3 and 4 and Figs 5 and 6 give the whole numerical results and another two sets of visual displays of HR images reconstructed by various methods. In contrast to their previous performance, it can be seen that the SPSR method is the most sensitive to the perturbation caused by noise (especially in scenario 4), which results in not only serious noise-caused artifacts in recovered images but also very



**Fig. 3** Visual comparison on *Leaves* in scenario 1 with scale factor 3. **a** Original image, HR images reconstructed by **b** SCSR [50], **c** SLSR [55], **d** SPSR [34], **e** ASDS [12], **f** NCSR [14], **g** ANR [39], **h** A+ [40], **i** SRCNN [15] and **j** proposed BASR

severe declines in numerical indicators. The reason for this is that a particular set of parameters in the SPSR model is trained to work efficiently under one certain condition. If the circumstances are slightly changed, all parameters need to be retrained or reassigned. In other words, the SPSR method lacks the characteristic of robustness to noise, which is the fatal drawback of it. Similarly, although both SCSR and SLSR methods can indisputably enhance the resolution of LR images, they are still liable to generate noticeable jaggy artifacts along image edges, indicating that using sparse prior only is not enough for reliable reconstructions when facing noise. As for ANR, A+ and SRCNN, they do much better in reducing jaggy artifacts, whereas they still magnify, rather than eliminate, the unpleasant influence of the miscellaneous outliers to some degree. The ASDS and NCSR methods both show an outstanding capability of recovering high frequency components and suppressing noise; nevertheless, they tend to produce some unclear and blurred local details or even incorrect parts sometimes. For instance, in Fig. 6(e) the shape of the left eye of *Lena* is kind of deformed, while a wrong white spot can



**Fig. 4** Visual comparison on *Parrot* in scenario 2 with scale factor 3. **a** Original image, HR images reconstructed by **b** SCSR [50], **c** SLSR [55], **d** SPSR [34], **e** ASDS [12], **f** NCSR [14], **g** ANR [39], **h** A+ [40], **i** SRCNN [15] and **j** proposed BASR

**Table 3** Numerical results (PSNR and SSIM) of the reconstructed HR images in scenario 3

Images	SCSR [50]	SLSR [55]	SPSR [34]	ASDS [12]	NCSR [14]	ANR [39]	A+ [40]	SRCNN [15]	BASR
Bike	23.37	23.69	24.10	24.22	24.44	23.75	24.33	24.35	<b>24.53</b>
	0.7246	0.7459	0.7629	0.7667	0.7788	0.7476	0.7712	0.7671	<b>0.7802</b>
Butterfly	24.56	25.68	26.16	26.62	27.56	25.56	26.66	27.26	<b>28.03</b>
	0.7872	0.8462	0.8610	0.8762	0.9032	0.8358	0.8665	0.8659	<b>0.9076</b>
Hat	29.24	29.76	30.02	30.21	30.59	29.59	30.13	29.99	<b>30.65</b>
	0.7696	0.7998	0.8079	0.8298	0.8392	0.7891	0.7952	0.7876	<b>0.8420</b>
Leaves	23.93	25.07	25.45	26.16	27.03	24.96	25.79	26.16	<b>27.75</b>
	0.8050	0.8553	0.8679	0.8869	0.9123	0.8470	0.8781	0.8838	<b>0.9214</b>
Parrot	28.23	28.90	29.21	29.74	30.00	28.93	29.33	29.58	<b>30.10</b>
	0.8205	0.8505	0.8560	0.8861	0.8863	0.8421	0.8429	0.8418	<b>0.8893</b>
Flower	27.49	28.15	28.33	28.59	28.62	28.20	28.52	28.54	<b>28.82</b>
	0.7690	0.7977	0.8056	0.8152	0.8217	0.7963	0.8056	0.8013	<b>0.8255</b>
Girl	31.58	32.17	32.17	32.51	32.49	32.06	31.98	31.92	<b>32.59</b>
	0.7569	0.7736	0.7760	0.7866	0.7857	0.7685	0.7637	0.7666	<b>0.7884</b>
Parthenon	25.91	26.45	26.49	26.67	26.76	26.32	26.59	26.70	<b>26.92</b>
	0.6730	0.6976	0.7048	0.7149	0.7216	0.6900	0.6981	0.7001	<b>0.7243</b>
Plants	30.65	31.48	31.59	32.05	32.51	31.39	31.84	31.84	<b>32.63</b>
	0.8194	0.8491	0.8521	0.8675	0.8826	0.8430	0.8481	0.8418	<b>0.8829</b>
Raccoon	28.07	28.43	28.50	28.64	28.66	28.43	28.55	28.65	<b>28.74</b>
	0.7153	0.7296	0.7298	0.7243	0.7282	0.7296	0.7311	0.7325	<b>0.7357</b>
Lena	29.74	30.55	30.79	31.12	31.29	30.52	30.88	30.87	<b>31.41</b>
	0.7908	0.8232	0.8266	0.8480	0.8536	0.8177	0.8212	0.8180	<b>0.8549</b>
Starfish	26.94	27.64	28.07	28.40	28.65	27.61	27.94	28.37	<b>28.68</b>
	0.7939	0.8199	0.8309	0.8415	0.8483	0.8167	0.8237	0.8321	<b>0.8492</b>
Zebra	25.02	26.46	26.45	27.13	27.27	26.28	26.91	27.26	<b>27.77</b>
	0.7979	0.8400	0.8425	0.8542	0.8558	0.8351	0.8474	0.8526	<b>0.8618</b>
Average	28.29	28.03	28.26	28.62	28.91	27.97	28.42	28.58	<b>29.12</b>
	0.7710	0.8022	0.8095	0.8229	0.8321	0.7968	0.8071	0.8070	<b>0.8356</b>

The numbers in bold indicate the best performance (i.e., the highest values)

be observed in Fig. 6(f), which all make the whole HR images reconstructed by ASDS and NCSR look unnatural and uncomfortable. On the contrary, by taking advantage of bidirectional similarities, our BASR method can produce more accurate sparse coding coefficients so that it not only suppresses the noise efficiently but also preserves more delicate structures as compared to any other counterparts. Thus, its corresponding HR outputs are still of the best visual quality as we can see, showing a strong robustness against the noise.

### 4.3 Comparison and analysis on the rate of convergence

In this subsection, we decide to make an investigation on the convergence rate of our method. The second scenario is selected as the typical testing environment, and be aware that not all the competitive methods are added to the comparison this time, since some of them are not iteratively solved. Consequently, the ASDS and NCSR methods are picked as the testing baselines. In addition to the two methods, a variant of NCSR model, in which the FISTA algorithm is employed to solve it in place of using its original solution, is also implemented for the purpose of testing the effectiveness of FISTA and giving a more comprehensive illustration. The variant will be referred to as “NCSR + FISTA” hereinafter.

Then, the RMSE values of the first five images (*bike*, *butterfly*, *hat*, *leaves* and *parrot*) varying with the increase of iterations are plotted in Fig. 7(a)–(e), respectively. According to

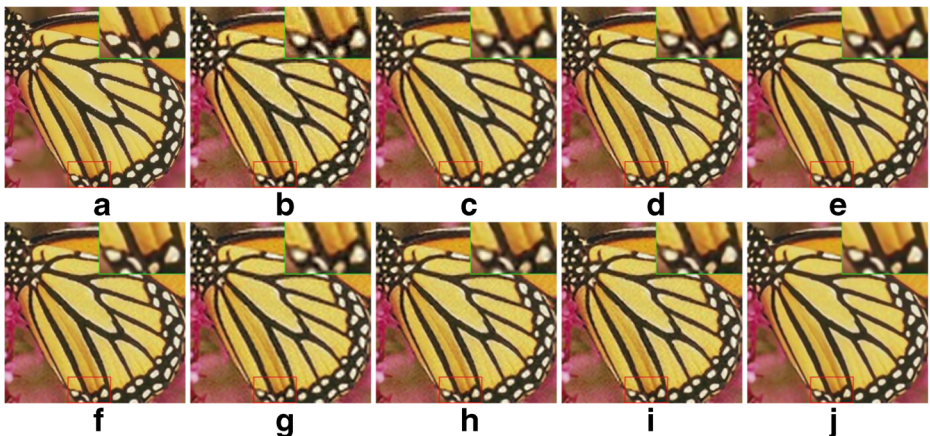


**Table 4** Numerical results (PSNR and SSIM) of the reconstructed HR images in scenario 4

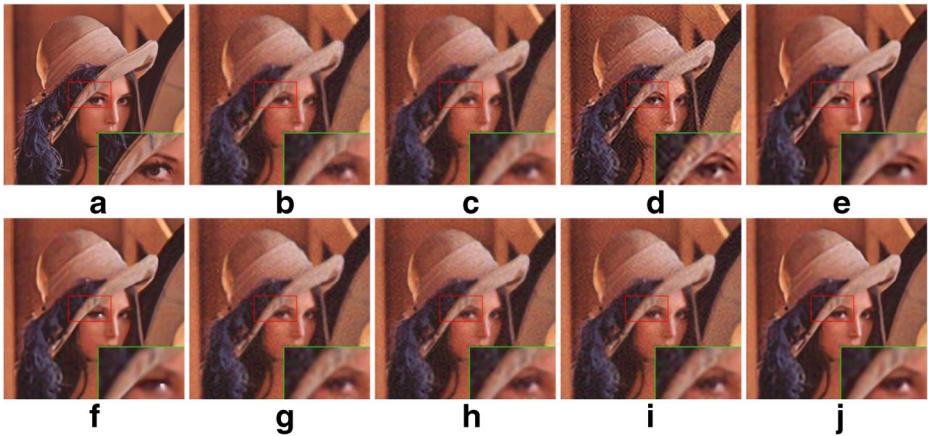
Images	SCSR [50]	SLSR [55]	SPSR [34]	ASDS [12]	NCSR [14]	ANR [39]	A+ [40]	SRCNN [15]	BASR
Bike	22.65	22.61	23.24	23.54	23.77	22.61	22.88	22.91	<b>24.01</b>
Butterfly	0.6688	0.6761	0.6860	0.7219	0.7361	0.6759	0.6881	0.6874	<b>0.7443</b>
Butterfly	23.74	23.84	24.99	25.91	26.69	23.71	24.16	24.22	<b>27.33</b>
Butterfly	0.7696	0.8039	0.7686	0.8560	0.8837	0.7937	0.8094	0.8061	<b>0.8873</b>
Hat	28.50	28.83	28.15	29.51	29.95	28.61	28.90	28.77	<b>29.96</b>
Hat	0.7473	0.7728	0.6627	0.8091	<b>0.8213</b>	0.7603	0.7609	0.7548	0.8205
Leaves	23.08	23.14	24.25	25.25	26.20	23.02	23.41	23.44	<b>26.96</b>
Leaves	0.7678	0.7938	0.7971	0.8613	0.8928	0.7852	0.8046	0.8040	<b>0.8982</b>
Parrot	27.52	27.53	27.45	29.03	29.45	27.54	27.76	27.75	<b>29.46</b>
Parrot	0.8017	0.8236	0.7042	0.8687	0.8711	0.8151	0.8127	0.8104	<b>0.8729</b>
Flower	26.88	26.94	26.77	27.80	27.95	26.94	27.10	27.11	<b>28.06</b>
Flower	0.7307	0.7447	0.6965	0.7776	0.7894	0.7420	0.7459	0.7422	<b>0.7928</b>
Girl	31.16	31.43	29.15	31.85	31.96	31.34	31.28	31.27	<b>32.02</b>
Girl	0.7301	0.7401	0.6344	0.7600	0.7624	0.7353	0.7293	0.7332	<b>0.7637</b>
Parthenon	25.55	25.74	25.45	26.23	26.44	25.66	25.82	25.85	<b>26.51</b>
Parthenon	0.6412	0.6569	0.6107	0.6903	0.7001	0.6507	0.6534	0.6537	<b>0.7007</b>
Plants	29.97	30.23	28.90	31.00	31.69	30.10	30.30	30.24	<b>31.80</b>
Plants	0.7922	0.8123	0.7042	0.8350	0.8550	0.8045	0.8044	0.7991	<b>0.8594</b>
Raccoon	27.58	27.67	26.83	28.02	28.02	27.67	27.70	27.79	<b>28.08</b>
Raccoon	0.6686	0.6753	0.6299	0.6810	0.6838	0.6752	0.6743	0.6760	<b>0.6925</b>
Lena	29.08	29.44	28.45	30.28	30.50	29.33	29.54	29.45	<b>30.62</b>
Lena	0.7661	0.7898	0.6875	0.8225	0.8308	0.7836	0.7828	0.7789	<b>0.8330</b>
Starfish	26.22	26.31	26.50	27.49	27.74	26.31	26.45	26.53	<b>27.81</b>
Starfish	0.7558	0.7685	0.7349	0.8057	0.8141	0.7662	0.7684	0.7706	<b>0.8157</b>
Zebra	24.16	24.24	25.24	26.33	26.59	24.13	24.46	24.62	<b>27.10</b>
Zebra	0.7553	0.7736	0.7765	0.8218	0.8290	0.7696	0.7773	0.7831	<b>0.8354</b>
Average	26.62	26.77	26.57	27.86	28.22	26.69	26.90	26.92	<b>28.44</b>
Average	0.7381	0.7563	0.6995	0.7931	0.8054	0.7506	0.7547	0.7538	<b>0.8090</b>

The numbers in bold indicate the best performance (i.e., the highest values)

the graphics, the discoveries here are twofold: First, it can be observed that our method consistently converges fastest among these methods, which demonstrates both the efficiency



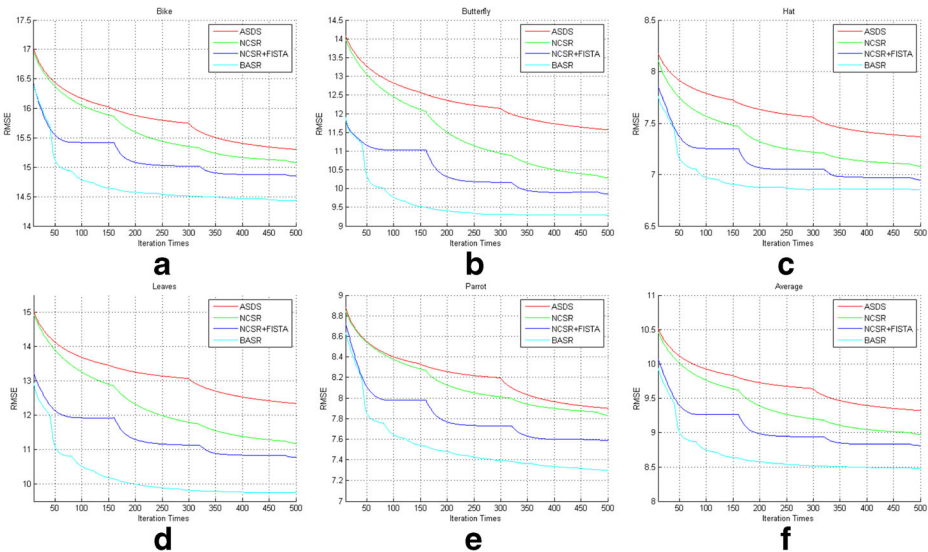
**Fig. 5** Visual comparison on *Butterfly* in scenario 3 with scale factor 3. **a** Original image, HR images reconstructed by **b** SCSR [50], **c** SLSR [55], **d** SPSR [34], **e** ASDS [12], **f** NCSR [14], **g** ANR [39], **h** A+ [40], **i** SRCNN [15] and **j** proposed BASR



**Fig. 6** Visual comparison on *Lena* in scenario 4 with scale factor 3. **a** Original image, HR images reconstructed by **b** SCSR [50], **c** SLSR [55], **d** SPSR [34], **e** ASDS [12], **f** NCSR [14], **g** ANR [39], **h** A+ [40], **i** SRCNN [15] and **j** proposed BASR

and effectiveness of the proposed BASR model. Second, we still found that the original NCSR method is inferior to its variant, NCSR + FISTA, in terms of convergence rate. The improvement on the variant can only be attributed to the virtue of using FISTA algorithm, since we exactly preserved the whole framework of NCSR (even the original parameter settings were maintained) except for the change on solution. And because we had not fine-tuned the parameters of NCSR + FISTA, it becomes reasonable that its curves are ladder-like and not smooth enough.

Furthermore, to provide more precise information, we finished the experiments on the remaining samples. The average RMSE curve of all images is plot in Fig. 7(f). It can be seen



**Fig. 7** Convergence rates of ASDS [12], NCSR [14], NCSR + FISTA and BASR methods. RMSE values on **(a)** *Bike*, **(b)** *Butterfly*, **(c)** *Hat*, **(d)** *Leaves*, **(e)** *Parrot* and **(f)** Average of all test samples

that on average ASDS and NCSR do not converge after 500 iterations (actually after 900, as not shown in the figure), whereas our method only needs 340 ~ 380 iterations to yield a convergence. This is also the reason that we set the number of iterations equal to 360 times. As for running time, the ASDS, NCSR and BASR methods cost about 298, 532, 319 s respectively, to reconstructed a super-resolved image from  $85 \times 85$  to  $255 \times 255$  pixels. Even though ASDS requires the least amount of time, this is acquired by virtue of offline learning and never updating dictionary. If its learning process runs once, it will cost another 10 min. By contrast, our method achieves equivalent performance on running time as compared to ASDS, while costs much less when compared with NCSR, showing the efficiency once again.

In conclusion, according to the conducted experiments, the FISTA algorithm can be acknowledged as a more promising solution to BASR in the sense that it has a higher rate of convergence, and the superiority of our method to similar ones is fully verified.

## 5 Conclusion

In this paper, a bidirectionally aligned sparse representation model was proposed for the application of single image super-resolution. Motivated by the recent discovery [57] that image patch similarities exist not only among columns but also among rows of similarity data matrix, the modeling of bidirectional similarities is first presented in our model. By virtue of it, we constructed a pair of regularization terms in the form of  $l_1$ -norm, to which the raw sparse coefficients are simultaneously aligned after sparse coding process in order to make up for the errors caused by image noise and degradation. On the basis of FISTA algorithm, a local-to-global and coarse-to-fine solving course was developed to efficiently solve the proposed BASR model. Extensive experiments were performed to make a complete comparison between the BASR method and other leading methods, and the results indicate that our method is constantly superior to its counterparts in terms of both numerical assessment and visual perception. In our future work, we will concentrate on the research on the computational efficiency and adaptive allocation of weights of similarities and regularization in order to achieve a further improvement.

**Acknowledgments** The authors would like to thank the associate editor and anonymous reviewers for their constructive and precious comments, which helped us a lot in improving the presentation of this work.

This work was supported by the National Natural Science Foundation of China (No.61374194, No.61403081), the National Key Science & Technology Pillar Program of China (No.2014BAG01B03), the Key Research and Development Program of Jiangsu Province (No. BE2016739), and a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

## Appendix A: Solving BASR on the basis of FISTA

Here, we provide the main process for solving the proposed BASR model (11) on the basis of FISTA [3] which is an effective algorithm for linear inverse problems with dense matrix data. Even though this part is more detailed when compared with that in subsection 3.4, it is still a sketch only. For comprehensive details, we direct you to [3, 11].

First of all, for the sake of convenience, let us divide (11) into two functions and define them respectively

$$\begin{aligned}
 f(\alpha_i) &\triangleq \|\mathbf{y}_i - \mathbf{H}\mathbf{D}_{i_i}\alpha_i\|_2^2, \quad \nabla f(\alpha_i) = 2\mathbf{D}_{i_i}^T(\mathbf{H}^T\mathbf{H}\mathbf{x}_i - \mathbf{H}^T\mathbf{y}_i) \\
 g(\alpha_i) &\triangleq \mu_1\|\alpha_i - \varphi_i\|_1 + \mu_2\|\alpha_i - \psi_i\|_1
 \end{aligned}
 \tag{13}$$

Then, according to the mathematic results derived in [3], model (11) can be approximated by considering the following formula at the given point  $\alpha_i^{(l)}$  (i.e., the sparse coefficient in the  $l$ -th iteration)

$$\begin{aligned}
 \{\alpha_i^{(l+1)}\}_{i=1}^Q &= \rho(\alpha_i^{(l)}, L) \triangleq \arg \min_{\alpha_i^{(l+1)}} \left\{ \omega(\alpha_i^{(l+1)}, \alpha_i^{(l)}, L) \right\} \\
 &= \arg \min_{\alpha_i^{(l+1)}} \left\{ f(\alpha_i^{(l)}) + \langle \alpha_i^{(l+1)} - \alpha_i^{(l)}, \nabla f(\alpha_i^{(l)}) \rangle + \frac{L}{2} \|\alpha_i^{(l+1)} - \alpha_i^{(l)}\|_2^2 + g(\alpha_i^{(l+1)}) \right\}, \\
 &\text{for } i = 1, 2, \dots, Q
 \end{aligned}
 \tag{14}$$

It has been proved that the above formula admits a unique minimizer. By getting rid of the constant terms, function  $\rho$  can be reformulated into a briefer from

$$\rho(\alpha_i^{(l)}, L) = \arg \min_{\alpha_i^{(l+1)}} \left\{ g(\alpha_i^{(l+1)}) + \frac{L}{2} \left\| \alpha_i^{(l+1)} - \left( \alpha_i^{(l)} - \frac{1}{L} \nabla f(\alpha_i^{(l)}) \right) \right\|_2^2 \right\}
 \tag{15}$$

Notation  $L$  here is involved to control the magnitude of step-size. After taking the derivative of (15) with respect to  $\alpha_i^{(l+1)}$  and making it equal to zero, we can get

$$\alpha_i^{(l+1)} + \frac{\mu_1}{L} \text{sgn}(\alpha_i^{(l+1)} - \varphi_i^{(l)}) + \frac{\mu_2}{L} \text{sgn}(\alpha_i^{(l+1)} - \psi_i^{(l)}) = \alpha_i^{(l)} - \frac{1}{L} \nabla f(\alpha_i^{(l)})
 \tag{16}$$

If we let the left side part of (16) form another new function  $F$ , that is

$$F(\alpha_i^{(l+1)}) = \alpha_i^{(l+1)} + \frac{\mu_1}{L} \text{sgn}(\alpha_i^{(l+1)} - \varphi_i^{(l)}) + \frac{\mu_2}{L} \text{sgn}(\alpha_i^{(l+1)} - \psi_i^{(l)})
 \tag{17}$$

Then, the function  $\rho$  can be expressed as the inverse function of  $F$

$$\rho(\alpha_i^{(l+1)}, L) = F^{-1} \left( \alpha_i^{(l)} - \frac{1}{L} \nabla f(\alpha_i^{(l)}) \right) = F^{-1} \left( \mathbf{D}_{i_i}^T \left( \mathbf{x}_i^{(l)} - \frac{2}{L} (\mathbf{H}^T \mathbf{H} \mathbf{x}_i^{(l)} - \mathbf{H}^T \mathbf{y}_i) \right) \right)
 \tag{18}$$

And it can be readily formulated as

If  $\varphi \leq \psi$ ,

$$\rho(\alpha, L) = \begin{cases} \alpha + \frac{\mu_1}{L} + \frac{\mu_2}{L} & \text{if } \alpha < \varphi - \frac{\mu_1}{L} - \frac{\mu_2}{L} \\ \varphi & \text{if } \varphi - \frac{\mu_1}{L} - \frac{\mu_2}{L} \leq \alpha < \varphi + \frac{\mu_1}{L} - \frac{\mu_2}{L} \\ \alpha - \frac{\mu_1}{L} + \frac{\mu_2}{L} & \text{if } \varphi + \frac{\mu_1}{L} - \frac{\mu_2}{L} \leq \alpha < \psi + \frac{\mu_1}{L} - \frac{\mu_2}{L} \\ \psi & \text{if } \psi + \frac{\mu_1}{L} - \frac{\mu_2}{L} \leq \alpha < \psi + \frac{\mu_1}{L} + \frac{\mu_2}{L} \\ \alpha - \frac{\mu_1}{L} - \frac{\mu_2}{L} & \text{if } \psi + \frac{\mu_1}{L} + \frac{\mu_2}{L} \leq \alpha \end{cases}
 \tag{19}$$

Else if  $\varphi > \psi$ ,

$$\rho(\alpha, L) = \begin{cases} \alpha + \frac{\mu_1}{L} + \frac{\mu_2}{L} & \text{if } \alpha < \psi - \frac{\mu_1}{L} - \frac{\mu_2}{L} \\ \psi & \text{if } \psi - \frac{\mu_1}{L} - \frac{\mu_2}{L} \leq \alpha < \psi - \frac{\mu_1}{L} + \frac{\mu_2}{L} \\ \alpha + \frac{\mu_1}{L} - \frac{\mu_2}{L} & \text{if } \psi - \frac{\mu_1}{L} + \frac{\mu_2}{L} \leq \alpha < \varphi - \frac{\mu_1}{L} + \frac{\mu_2}{L} \\ \varphi & \text{if } \varphi - \frac{\mu_1}{L} + \frac{\mu_2}{L} \leq \alpha < \varphi + \frac{\mu_1}{L} + \frac{\mu_2}{L} \\ \alpha - \frac{\mu_1}{L} - \frac{\mu_2}{L} & \text{if } \varphi + \frac{\mu_1}{L} + \frac{\mu_2}{L} \leq \alpha \end{cases} \quad (20)$$

It can be noted that the above algorithm corresponds to the general process of classic iterative thresholding algorithm (ITA) [11], which can be viewed as an extension of the classic gradient-based method. Function  $\rho$  acts as the shrinkage operator and processed in a pixelwise manner. Although it is well known that the first order optimization algorithms are often the only simple and practical option to deal with large-scale problems such as the case in this paper, it still has been found that the sequence produced by the above algorithm converges quite slowly to the final minimizer [4]. To accelerate the algorithm, the temporary variable before shrinkage operation in (18) ought not to be computed by considering the result obtained in the previous iteration only, but rather to be calculated by utilizing a very special linear combination of the previous two results. Therefore, formula (18) can be modified as follows:

$$\begin{cases} \mathbf{Z} = \mathbf{X}^{(l)} + \frac{t^{(l-1)} - 1}{t^{(l)}} (\mathbf{X}^{(l)} - \mathbf{X}^{(l-1)}) \\ \rho(\boldsymbol{\alpha}_i^{tmp}, L) = \rho\left(\mathbf{D}_i^T \mathbf{R}_i \left(\mathbf{Z} - \frac{2}{L} (\mathbf{H}^T \mathbf{H} \mathbf{Z} - \mathbf{H}^T \mathbf{Y})\right), L\right) \end{cases} \quad (21)$$

Besides, another condition ensuring convergence is to require that the step-size controller  $L$  is set to be no less than the smallest Lipschitz constant of the gradient of function  $f$ , and this quantity can be determined by employing a backtracking step-size rule. In summary, a step-by-step description of the above implementation can be given in Algorithm 1.

## References

1. Aguen ML, Mascarenhas ND (2006) Multispectral image data fusion using POCS and super-resolution. *Comput Vis Image Underst* 102(2):178–187
2. Baker S, Kanade T (2002) Limits on super-resolution and how to break them. *IEEE Trans Pattern Anal Mach Intell* 24(9):1167–1183
3. Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imag Sci* 2(1):183–202
4. Bioucas-Dias JM, Figueiredo MAT (2007) A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans Image Process* 16(12):2992–3004

5. Bose NK, Kim HC, Valenzuela HM (1993) Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframe. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 269–272
6. Buades A, Coll B, Morel J (2005) A non-local algorithm for image denoising. In: IEEE computer society conference on Computer Vision and Pattern Recognition (CVPR), 60–65
7. Candès EJ, Wakin MB (2008) An introduction to compressive sampling. *IEEE Signal Process Mag* 25(2):21–30
8. Candès EJ, Romberg JK, Tao T (2006) Stable signal recovery from incomplete and inaccurate measurements. *Commun Pure Appl Math* 59(8):1207–1223
9. Candès EJ, Wakin MB, Boyd SP (2008) Enhancing sparsity by reweighted  $\ell_1$  minimization. *J Fourier Anal Appl* 14(5–6):877–905
10. Chang H, Yeung DY, Xiong Y (2004) Super-resolution through neighbor embedding. In: IEEE computer society conference on Computer Vision and Pattern Recognition (CVPR), I–I
11. Daubechies I, Defrise M, De Mol C (2004) An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun Pure Appl Math* 57(11):1413–1457
12. Dong W, Zhang L, Shi G, Wu X (2011) Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization. *IEEE Trans Image Process* 20(7):1838–1857
13. Dong W, Zhang L, Shi G (2011) Centralized sparse representation for image restoration. In: IEEE International Conference on Computer Vision (ICCV), 1259–1266
14. Dong W, Zhang L, Shi G, Li X (2013) Nonlocally centralized sparse representation for image restoration. *IEEE Trans Image Process* 22(4):1620–1630
15. Dong C, Loy CC, He K, Tang X (2016) Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Mach Intell* 38(2):295–307
16. Donoho DL (2006) For most large underdetermined Systems of Linear Equations the minimal 1-norm solution is also the sparsest solution. *Commun Pure Appl Math* 59(6):797–829
17. Elad M, Aharon M (2006) Image Denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans Image Process* 15(12):3736–3745
18. Elad M, Feuer A (1997) Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images. *IEEE Trans Image Process* 6(12):1646–1658
19. Elad M, Yavneh I (2009) A plurality of sparse representations is better than the sparsest one alone. *IEEE Trans Inf Theory* 55(10):4701–4714
20. Farsiu S, Robinson MD, Elad M, Milanfar P (2004) Fast and robust multiframe super resolution. *IEEE Trans Image Process* 13(10):1327–1344
21. Farsiu S, Robinson D, Elad M, Milanfar P (2004) Advances and challenges in super-resolution. *Int J Imaging Syst Technol* 14(2):47–57
22. Freeman WT, Jones TR, Pasztor EC (2002) Example-based super-resolution. *IEEE Comput Graph Appl* 22(2):56–65
23. He Y, Yap K, Chen L, Chau L (2009) A soft MAP framework for blind super-resolution image reconstruction. *Image Vis Comput* 27(4):364–373
24. Kim SP, Su W (1993) Recursive high-resolution reconstruction of blurred multiframe images. *IEEE Trans Image Process* 2(4):534–539
25. Kim SP, Bose NK, Valenzuela HM (1990) Recursive reconstruction of high resolution image from noisy undersampled multiframe. *IEEE Trans Acoust Speech Signal Process* 38(6):1013–1027
26. Li X, Hu Y, Gao X, Tao D, Ning B (2010) A multi-frame image super-resolution method. *Signal Process* 90(2):405–414
27. Lin ZC, Shum HY (2004) Fundamental limits of reconstruction-based superresolution algorithms under local translation. *IEEE Trans Pattern Anal Mach Intell* 26(1):83–97
28. Lu X, Yuan H, Yan P, Yuan Y, Li X (2012) Geometry constrained sparse coding for single image super-resolution. In: IEEE conference on Computer Vision and Pattern Recognition (CVPR), 1648–1655
29. Mishra D, Majhi B, Sa PK, Dash R (2016) Development of robust neighbor embedding based super-resolution scheme. *Neurocomputing* 202:49–66
30. Nasrollahi K, Moeslund TB (2014) Super-resolution: a comprehensive survey. *Mach Vis Appl* 25(6):1423–1468
31. Nguyen N, Milanfar P, Golub G (2001) A computationally efficient superresolution image reconstruction algorithm. *IEEE Trans Image Process* 10(4):573–583

32. Omer OA, Tanaka T (2011) Region-based weighted-norm with adaptive regularization for resolution enhancement. *Digital Signal Process* 21(4):508–516
33. Park SC, Park MK, Kang MG (2003) Super-resolution image reconstruction: a technical overview. *IEEE Signal Process Mag* 20(3):21–36
34. Peleg T, Elad M (2014) A statistical prediction model based on sparse representations for single image super-resolution. *IEEE Trans Image Process* 23(6):2569–2582
35. Purkait P, Pal NR, Chanda B (2014) A fuzzy-rule-based approach for single frame super resolution. *IEEE Trans Image Process* 23(5):2277–2290
36. Qin F, He X, Chen W, Yang X, Wu W (2009) Video superresolution reconstruction based on subpixel registration and iterative back projection. *J Electron Imaging* 18(1):13007
37. Stark H, Oskoui P (1989) High-resolution image recovery from image-plane arrays, using convex projections. *J Opt Soc Am A* 6(11):1715–1726
38. Tian J, Ma K (2011) A survey on super-resolution imaging. *SIViP* 5(3):329–342
39. Timofte R, De Smet V, Van Gool L (2013) Anchored neighborhood regression for fast example-based super-resolution. In: *IEEE International Conference on Computer Vision (ICCV)*, 1920–1927
40. Timofte R, De Smet V, Van Gool L (2014) A+: adjusted anchored neighborhood regression for fast super-resolution. In: *Asian Conference on Computer Vision (ACCV)*, 111–126
41. Timofte R, Rothe R, Van Gool L (2016) Seven ways to improve example-based single image super resolution. In: *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, 1865–1873
42. Tsai RY, Huang TS (1984) Multiframe image restoration and registration. *Advances in computer vision and Image Processing* 1(2):317–339
43. Vrigkas M, Nikou C, Kondi LP (2013) Accurate image registration for MAP image super-resolution. *Signal Process Image Commun* 28(5):494–508
44. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process* 13(4):600–612
45. Yan C, Zhang Y, Dai F, Wang X, Li L, Dai Q (2014) Parallel deblocking filter for HEVC on many-core processor. *Electron Lett* 50(5):367–368
46. Yan C, Zhang Y, Dai F, Zhang J, Li L, Dai Q (2014) Efficient parallel HEVC intra-prediction on many-core processor. *Electron Lett* 50(11):805–806
47. Yan C, Zhang Y, Xu J, Dai F, Li L, Dai Q, Wu F (2014) A highly parallel framework for HEVC coding unit partitioning tree decision on many-core processors. *IEEE Signal Process Lett* 21(5):573–576
48. Yan C, Zhang Y, Xu J, Dai F, Zhang J, Dai Q, Wu F (2014) Efficient parallel framework for HEVC motion estimation on many-core processors. *IEEE Trans Circuits Syst Video Technol* 24(12):2077–2089
49. Yang J, Wright J, Huang T, Ma Y (2008) Image super-resolution as sparse representation of raw image patches. In: *IEEE conference on Computer Vision and Pattern Recognition (CVPR)*, 1–8
50. Yang J, Wright J, Huang TS, Ma Y (2010) Image super-resolution via sparse representation. *IEEE Trans Image Process* 19(11):2861–2873
51. Yang S, Liu Z, Wang M, Sun F, Jiao L (2011) Multitask dictionary learning and sparse representation based single-image super-resolution reconstruction. *Neurocomputing* 74(17):3193–3203
52. Yue L, Shen H, Yuan Q, Zhang L (2014) A locally adaptive  $L_1$ - $L_2$  norm for multi-frame super-resolution of images with mixed noise and outliers. *Signal Process* 105:156–174
53. Zeng W, Lu X (2012) A generalized DAMRF image modeling for superresolution of license plates. *IEEE Trans Intell Transp Syst* 13(2):828–837
54. Zeng W, Lu X, Fei S (2015) Image super-resolution employing a spatial adaptive prior model. *Neurocomputing* 162:218–233
55. Zeyde R, Elad M, Protter M (2012) On single image scale-up using sparse-representations. *Curves and Surfaces (Springer)*, 711–730.
56. Zhang X, Wu X (2008) Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation. *IEEE Trans Image Process* 17(6):887–896
57. Zhang X, Feng X, Wang W (2013) Two-direction nonlocal model for image Denoising. *IEEE Trans Image Process* 22(1):408–412
58. Zhou L, Lu X, Yang L (2014) A local structure adaptive super-resolution reconstruction method based on BTV regularization. *Multimedia Tools and Applications* 71(3):1879–1892



**Chao Xie** received the B.S. degree in automation and the M.S. degree in control theory and control engineering from Nanjing Forestry University, Nanjing, China, in 2010 and 2013, respectively. He is currently working toward the Ph.D. degree with the School of Automation, Southeast University. His current research interests include image processing and sparse signal representation.



**Weili Zeng** received the B. S. degree in Mathematics and Computer Science from Jishou University, Jishou, China, in 2007, the M.S. degree in Applied Mathematics and the Ph. D. degree in traffic image from Southeast University, Nanjing, China, in 2009 and 2013, respectively. He is a vice professor at College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include image processing, mathematical inverse problems. Dr. Zeng is a Reviewer of the IEEE Transactions on Intelligent Transportation System, IET Image Processing, and the Imaging Science Journal.





**Shengqin Jiang** received the M.S. degree in basic mathematics from Jiangsu University, Zhenjiang, China, in 2015. He is currently pursuing the Ph.D. degree with the school of Automation, Southeast University. His current research interests include object detection and tracking, action recognition and synchronization of complex networks. He was a recipient of the Outstanding Master's Degree Thesis Award from Jiangsu Province, China, in 2016.



**Xiaobo Lu** received the B.S. degree from Shanghai Jiao Tong University, Shanghai, China, the M.S. degree from Southeast University, Nanjing, China, and the Ph.D. degree from Nanjing University of Aeronautics and Astronautics. He did his postdoctoral research with Chien-Shiung Wu Laboratory, Southeast University, from 1998 to 2000. He is currently a Professor with the School of Automation and deputy Director of the Detection Technology and Automation Research Institute, Southeast University. He is a coauthor of the book *An Introduction to the Intelligent Transportation Systems* (Beijing, China Communications, 2008). His research interests include image processing, signal processing, pattern recognition, and computer vision. Dr. Lu has received many research awards, such as the First Prize in Natural Science Award from the Ministry of Education of China and the prize in the Science and Technology Award of Jiangsu province.