

# Iterative projection reconstruction for fast and efficient image upsampling



Yang Zhao<sup>a,b</sup>, Rong-Gang Wang<sup>a,\*</sup>, Wei Jia<sup>b</sup>, Wen-Min Wang<sup>a</sup>, Wen Gao<sup>a</sup>

<sup>a</sup> School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, Shenzhen 518055, China

<sup>b</sup> School of Computer and Information, Hefei University of Technology, Hefei 230009, China

## ARTICLE INFO

Communicated by Gaofeng Meng

**Keywords:**

Upsampling

Single-image super-resolution

Projection reconstruction

## ABSTRACT

With the development of ultra-high-resolution display devices, the visual perception of fine texture details is becoming increasingly important. Traditional image upsampling methods suffer from either loss of high-frequency texture details or very high time cost. In this paper, we propose an iterative projection reconstruction (IPR) method for fast and efficient image upsampling. The proposed method refines high-frequency texture details with an iterative projection process, and utilizes the pre-computed projection matrix to accelerate the example-based image reconstruction. As a result, the proposed method can reproduce fine texture details with low time cost. Experimental results demonstrate that the proposed method outperforms some state-of-the-art methods.

## 1. Introduction

Image upsampling, which aims to recover a high-resolution image (HRI) from a low-resolution image (LRI), is a basic problem for various applications in machine vision and image processing, such as digital photographs editing, ultra-high-definition (UHD) display, medical image processing, and object recognition. Recently, some specific super-resolution scenarios have also drawn significant attention, such as face hallucination [44,45], depth-image upsampling [50–52]. It is an ill-posed problem since much information is lost during the down-sampling of original HRIs. How to reconstruct the missing details of high quality HRIs with low cost is still a challenging task.

One fundamental technique of image upsampling is the interpolation-based algorithm, such as nearest neighbor, bilinear, bicubic, and splines [1,2]. These kernel-based interpolations are efficient and fast. However, these traditional methods have two obvious demerits: firstly, these methods tend to produce some unnatural artifacts such as blurring and zigzag edges; Secondly, the simple interpolations cannot reproduce the lost high-frequency (HF) details. Recent interpolation-based methods try to suppress unnatural artifacts via different techniques or prior knowledge, such as edge property prior [3–5], different shapes of interpolated-grids [6–9], and edge sharpening process [10–12], etc. Unfortunately, these recent methods are able to refine sharp edges, but still cannot recover fine texture details.

Another classic type of image upsampling method is reconstruction-based, which imposes a basic similarity constraint that the down-sampling of reconstructed HRI result should be consistent with the original LRI. Early algorithms [13,14] utilize multi-frames to recon-

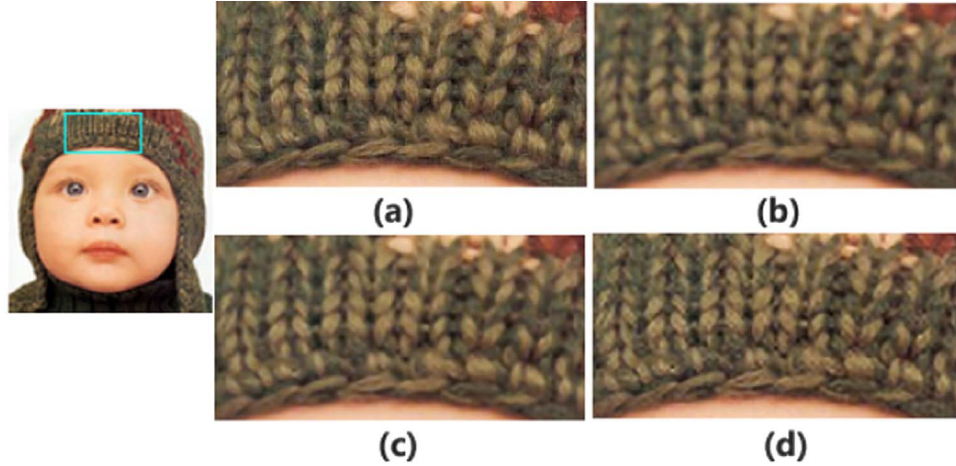
struct HRI, while more and more single-frame reconstruction-based upsampling methods are proposed recently to avoid the alignment problem. Recent methods often improve the basic similarity constraint with extra constraints or different models, such as, gradient prior constraint [15–19], local texture-structure constraint [47], steering kernel regression method [44] and some de-convolution based models [20–22], etc. However, the performance of reconstruction-based algorithms degrades rapidly when the upsampling factor becomes large, since the basic similarity is constrained at the LRI level.

To obtain the missing image details, many example-based or learning-based methods have been proposed over the years [23–42]. This type of method was first proposed in [23] by learning the missing information from extra low-resolution (LR) and high-resolution (HR) exemplars. In [25], Chang et al. proposed an effective learning method based on locally linear embedding. Yang et al. [26] presented an infusive sparse coding based method. Many sparse representation based methods were then proposed and achieved impressive results [27,28,31,34,35]. There are also some other example-based methods, such as Gaussian process regression based methods [54], combined example-based and reconstruction-based processes [53], special texture examples [46], and local self-exemplar based methods [38–40], etc. In recent years, many state-of-the-art methods have been proposed by means of deep neural network [55,56] and filter network [48]. Recent example-based methods can recover sharp and clear edges, but the improvement of texture details with finite exemplars/atoms is still a challenge work. Besides, the computational cost of traditional example-based methods is often quite high.

In order to reduce the time-cost of traditional dictionary-based

\* Corresponding author.

E-mail address: [rgwang@pkusz.edu.cn](mailto:rgwang@pkusz.edu.cn) (R.-G. Wang).



**Fig. 1.** Upsampled results of woolen texture with different methods (2× magnification). (a) Original HRI, (b) with bicubic interpolation, (c) with the ANR, (d) with the proposed IPR.

reconstruction, Timofte et al. [33,36] proposed fast anchored neighborhood regression methods (ANR), which simply estimate LR input by means of its similar dictionary atom and pre-compute the corresponding embedding matrix by utilizing the ridge regression. Comparing with the traditional example-based methods, the ANR can obtain one or two orders of magnitude speed-up. In [41], the pre-computed projection matrix is also applied to a sub-dictionaries-based method. In these ANR-based methods, the projection matrix is pre-computed and therefore the computation cost is not affected by the increase of neighbor samples. Hence, these methods have the potential capacity to refine the results with more neighbor exemplars. For example, the A+ method [36] pre-computes each projection matrix with more neighbor samples and can recover better edges than original ANR.

Unfortunately, the ANR (or A+) method still cannot obtain fine HF texture details. As illustrated in Fig. 1(c), the ANR reproduces blurring and unnatural HF texture details. That is mainly because the projection matrix of a LR input is estimated by directly utilizing the pre-computed projection matrix of its nearest atom. The accuracy of the reconstruction thus depends on the similarity between the LR input and its nearest dictionary atom, and the residual between the selected atom and the input is totally omitted in the ANR. This missing residual may lead to some reconstruction errors. As we known, the pre-learned dictionary often consists of stable edge-like atoms. However, the local texture patches usually contain abundant and various HF details. Hence, it is inaccurate to approximate a local texture patch by merely utilizing its nearest atom.

In this paper, we propose an iterative projection reconstruction (IPR) method for fast and efficient image upsampling. The proposed method recovers HF textural details with an iterative projection process, and utilizes the pre-computed projection matrix to accelerate the learning-based reconstruction in each iteration. In the proposed IPR, the residual components between the input patch and its nearest atom, which are omitted in the ANR, are also reconstructed by means of an iterative process. By comparing the textural details in Fig. 1(c) and (d), it can be found that the proposed IPR can recover finer texture details than the ANR. More experimental results also demonstrate that the proposed method can refine natural texture details with low time cost and outperform some state-of-the-art methods.

The following paragraphs of this paper are organized as follows. Section 2 briefly introduces related typical learning-based super-resolution models, and Section 3 presents the proposed method in detail. Section 4 gives the experimental results to testify the effectiveness of the proposed method, and Section 5 concludes the paper.

## 2. Related learning-based super-resolution models

In this section, we introduce several recent learning-based models, i.e., locally linear embedding (LLE), sparse representation, and the ANR.

### 2.1. Locally linear embedding based approaches

Chang et al. [25] proposed a super-resolution method based on neighbor embedding (NE) and the manifold learning method LLE. The NE method is based on the assumption that small patches in the LRIs and HRIs form manifold with similar local geometry in two distinct spaces. As long as enough samples are available, patches in the HR feature domain can be estimated as a weighted combination of local neighbors using the same weights as in the LR feature domain.

In example-based methods, HR dictionary  $\mathbf{D}_h$  and corresponding LR dictionary  $\mathbf{D}_l$  are pre-learned from training HRIs. For each input LR patch  $\mathbf{y}$ , Chang et al. search the set of  $K$  nearest neighbor atoms in the  $\mathbf{D}_l$  and compute the reconstruction weights of the neighbors  $\alpha_{LLE}$  that minimize the reconstruction error, as follows,

$$\min_{\alpha_{LLE}} \|\mathbf{y} - \mathbf{D}_l \alpha_{LLE}\|^2 \quad s. t. \quad \sum_{k=1}^K \alpha_{LLEk} = 1 \quad (1)$$

where  $\alpha_{LLE}$  only contains  $K$  weights for  $K$  nearest neighbor atoms and all the weights of other atoms are 0. The HR patch  $\mathbf{x}$  is then estimated with the reconstruction weights  $\alpha_{LLE}$  and  $\mathbf{D}_h$  by

$$\mathbf{x} = \mathbf{D}_h \alpha_{LLE} \quad (2)$$

### 2.2. Sparse representation based approaches

Yang et al. [26] proposed an effective sparse coding based super-resolution (ScSR) method. In the ScSR, coupled LR and HR dictionaries are jointly trained so that the LR patches and their corresponding HR patches can be represented with one sparse representation. The sparse representation of input LR patch  $\mathbf{y}$  is then calculated as a combination of dictionary atoms, as follows,

$$\min_{\alpha_{SC}} \|\mathbf{y} - \mathbf{D}_l \alpha_{SC}\|_2^2 + \lambda \|\alpha_{SC}\|_0 \quad (3)$$

where  $\alpha_{SC}$  denotes the representation coefficients, and  $\lambda$  is a weighting factor to balance the importance of the sparsity constraint. The  $l_0$ -norm constraint leads to an NP-hard problem and, in practice, is approximated by an  $l_1$ -norm constraint. After the representation coefficients  $\alpha_{SC}$  is calculated, the HR patch  $\mathbf{x}$  can be reconstructed in the same way as in Eq. (2).

### 2.3. Anchored neighborhood regression

The most time-consuming process in the LLE and the ScSR is calculating the weights or representation coefficients for each input LR patch. Timofte et al. [33] proposed a fast example-based method of ANR by pre-calculating the projection matrix of each atom and approximating the projection matrix of input patch with that of nearest atom.

In the ANR, the ridge regression is used to calculate the representation of LR input feature  $\mathbf{y}$ , as follows,

$$\min_{\beta} \|\mathbf{y} - N_i \beta\|_2^2 + \lambda \|\beta\|_2^2 \quad (4)$$

where  $N_i$  denotes the neighborhood in LR space, and  $\lambda$  is a weighting factor to stabilize the solution. The algebraic solution of Eq. (4) can be simply calculated by

$$\beta = (N_i^T N_i + \lambda I)^{-1} N_i^T \mathbf{y} \quad (5)$$

As a result, the HR patch can be estimated by means of the same coefficients:

$$\mathbf{x} = N_h \beta = N_h (N_i^T N_i + \lambda I)^{-1} N_i^T \mathbf{y} \quad (6)$$

where  $N_h$  is the HR neighborhood corresponding to  $N_i$ . The projection matrix is then defined by

$$\mathbf{P} = N_h (N_i^T N_i + \lambda I)^{-1} N_i^T \quad (7)$$

For a given LR input  $\mathbf{y}$ , the dictionary atom  $\mathbf{d}_i$  which is the most similar to  $\mathbf{y}$ , is first searched. The projection matrix of atom  $\mathbf{d}_i$  is then directly used as the projection matrix of  $\mathbf{y}$ . Finally, the HR patch is reconstructed with the estimated projection matrix  $\mathbf{P}_i$ :

$$\mathbf{x} = \mathbf{P}_i \mathbf{y} \approx \mathbf{P}_i \mathbf{y} \quad (8)$$

where  $\mathbf{P}_i$  is the projection matrix of the atom which is the most similar to  $\mathbf{y}$ . The projection matrix of each atom can be pre-computed offline by means of Eq. (7). Hence, the ANR can greatly save the time of reconstruction.

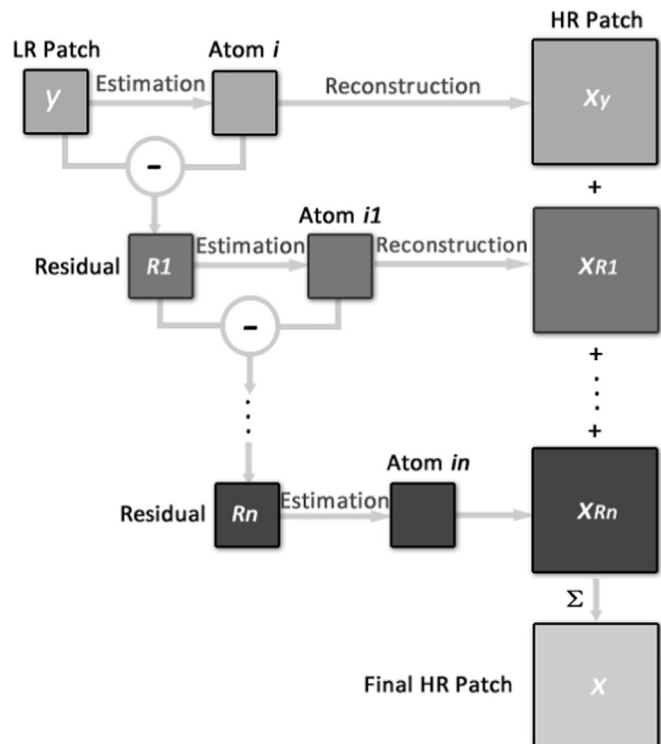


Fig. 2. Illustration of the iterative reconstruction process.

### 3. The proposed method

#### 3.1. Iterative projection reconstruction (IPR)

In the proposed iterative projection reconstruction process, we start also by grouping the dictionary atoms or training samples into neighborhoods as in [33,36]. For each dictionary atom we search its  $K$  nearest neighbors to set up its neighborhood. The  $K$  nearest neighbors of atom are selected with the use of Euclidean distance as the similarity measure. Each projection matrix of dictionary atom is then pre-calculated with Eq. (7). Hence, the proposed IPR inherits the high speed ability of ANR by means of pre-computed projection matrix.

Then, for each LR input  $\mathbf{y}$ , we first search its nearest neighbor atom as follows,

$$|\langle \mathbf{y}, \mathbf{d}_{q_0} \rangle| = \max_{n \in \{1, 2, \dots, N\}} |\langle \mathbf{y}, \mathbf{d}_n \rangle| \quad (9)$$

where  $\langle \cdot \rangle$  is the inner product,  $\max$  denotes the maximum,  $q_0$  is the index of the nearest dictionary atom, and  $N$  denotes the total number of the atoms. After the nearest atom  $\mathbf{d}_{q_0}$  is found, the corresponding projection matrix  $\mathbf{P}_{d_{q_0}}$  is used to estimate the projection matrix  $\mathbf{P}_y$  of input  $\mathbf{y}$ . The HR patch  $\mathbf{x}$  is then reconstructed by

$$\mathbf{x} = \mathbf{P}_y \mathbf{y} \approx \mathbf{P}_{d_{q_0}} \mathbf{y} \quad (10)$$

It should be noted that, in practice, all the training patches and LR patches have been turned into vector form and have been removed the mean-value. All the learned atoms are then normalized to represent a set of bases in residual space. Hence, the inner product is utilized to search the nearest atom on which the input has the longest projective length. At last, the reconstructed HR vector is turned to HR residual patch.

The missing residual between  $\mathbf{y}$  and its nearest atom  $\mathbf{d}_{q_0}$  is,

$$\mathbf{R}_1 = \mathbf{y} - \langle \mathbf{y}, \mathbf{d}_{q_0} \rangle \mathbf{d}_{q_0} \quad (11)$$

We then apply one more iteration to reconstruct the HR components of the missing residual  $\mathbf{R}_1$ . The second iteration is the repeat of the first iteration. The nearest neighbor atom of the residual  $\mathbf{R}_1$  is searched as follows

$$|\langle \mathbf{R}_1, \mathbf{d}_{q_1} \rangle| = \max_{n \in \{1, 2, \dots, N\}} |\langle \mathbf{R}_1, \mathbf{d}_n \rangle| \quad (12)$$

The projection matrix  $\mathbf{P}_{d_{q_1}}$  of the nearest atom  $\mathbf{d}_{q_1}$  is then used as the projection matrix of the residual  $\mathbf{R}_1$ . As a result, the reconstructed HR residual component is calculated by

$$\mathbf{x}_{R_1} = \mathbf{P}_{R_1} \mathbf{R}_1 \approx \mathbf{P}_{d_{q_1}} \mathbf{R}_1 \quad (13)$$

Then, the HR output patch  $\mathbf{x}$  is reconstructed by combining the reconstruction of  $\mathbf{y}$  and the reconstructed residual as follows,

$$\mathbf{x} = \mathbf{x}_y + \omega_1 \mathbf{x}_{R_1} = \mathbf{P}_{d_{q_0}} \mathbf{y} + \omega_1 \mathbf{P}_{d_{q_1}} \mathbf{R}_1 \quad (14)$$

where  $\omega_1$  ( $\omega_1 < 1$ ) is a weight to control the contribution of the residual term.

Similarly, the proposed method can be simply extended to more iterations. For example, the  $i$ -th residual is computed as,

$$\mathbf{R}_i = \mathbf{R}_{i-1} - \langle \mathbf{R}_{i-1}, \mathbf{d}_{q_{i-1}} \rangle \mathbf{d}_{q_{i-1}} \quad (15)$$

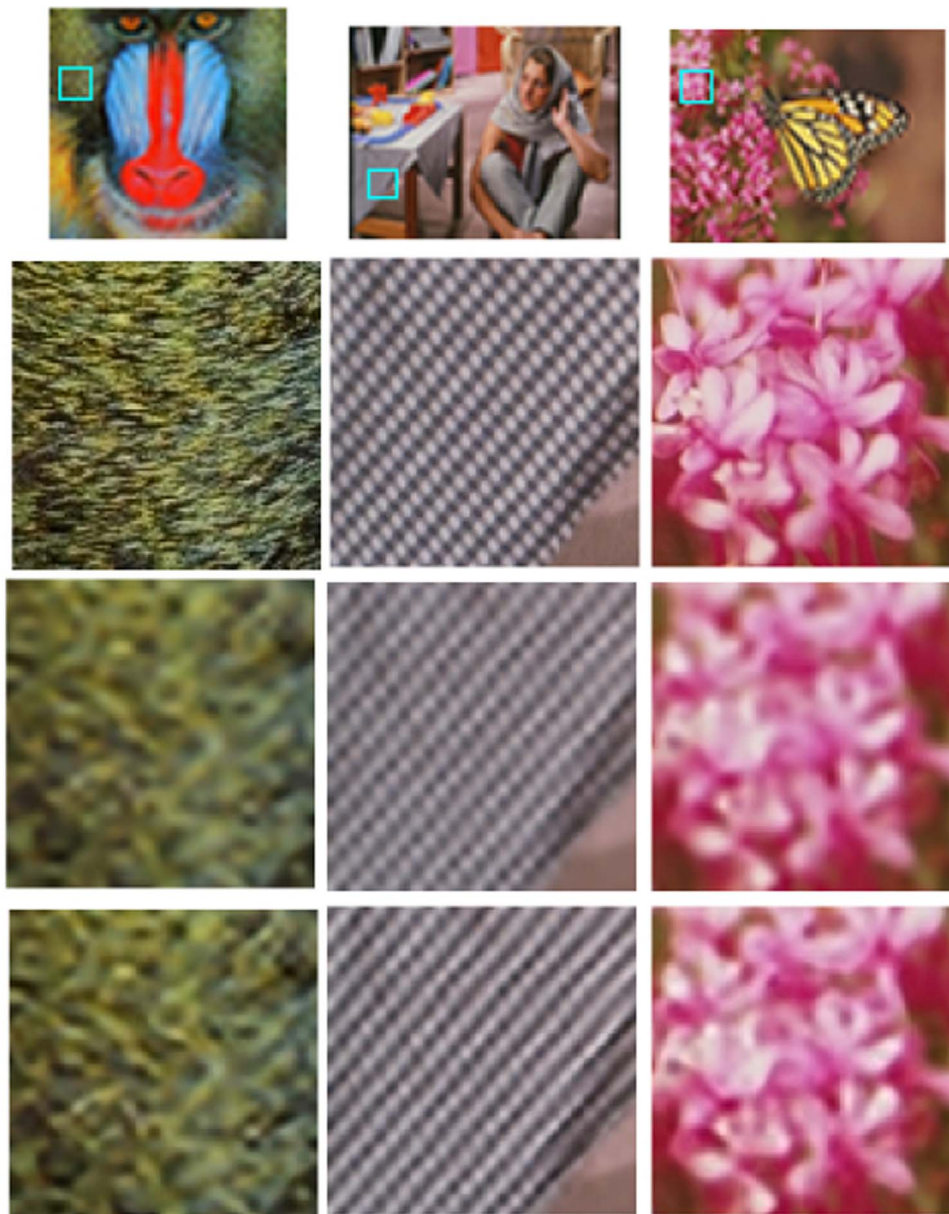
Its nearest neighbor atom  $\mathbf{d}_{q_i}$  is then selected by calculating the max inner product. After the projection matrix of each residual is estimated, the final HR output can be calculated by

$$\mathbf{x} = \mathbf{x}_y + \sum_{i=1}^{N_R} \omega_i \mathbf{x}_{R_i} = \mathbf{P}_{d_{q_0}} \mathbf{y} + \sum_{i=1}^{N_R} \omega_i \mathbf{P}_{d_{q_i}} \mathbf{R}_i \quad (16)$$

where  $N_R$  is the total number of iterations.

In this paper, we use the weight  $\omega_i$  to control the contribution of the HF residuals.<sup>1</sup> The weight is defined according to the following

<sup>1</sup> Related experimental results of different kinds of weights can be obtained from the



**Fig. 3.** Illustration of upsampling results with the ANR and the proposed IPR (4X magnification). The first row shows the selected texture areas; the second row shows the ground truth; the third row shows the close-ups upsampling with the ANR; the last row shows the close-ups upsampling with the IPR.

properties. Firstly, the weight  $\omega_i$  should be consistent with the magnitude of residual. It is obviously that the larger weight should be used for the larger residual. Secondly, the weight should also be consistent with the relative magnitude of residual by comparing with the former residual. Lastly, in order to make sure the strict convergence of the IPR process, the weight should be monotonically decreased. As a result, the weight  $\omega_i$  is simply defined as,

$$\omega_i = \begin{cases} \frac{\|R_i\|_2}{\|R_{i-1}\|_2 + \epsilon}, & \text{if } \frac{\|R_i\|_2}{\|R_{i-1}\|_2 + \epsilon} < \omega_{i-1} \\ \omega_{i-1}, & \text{if } \frac{\|R_i\|_2}{\|R_{i-1}\|_2 + \epsilon} \geq \omega_{i-1} \end{cases} \quad (17)$$

where  $i \in (1, 2, \dots, N_R)$ ,  $R_0 = y$ ,  $\omega_0 = 1$ , and  $\epsilon$  is a small positive number to prevent denominator from being zero (The convergence of the proposed IPR is explained in Appendix A).

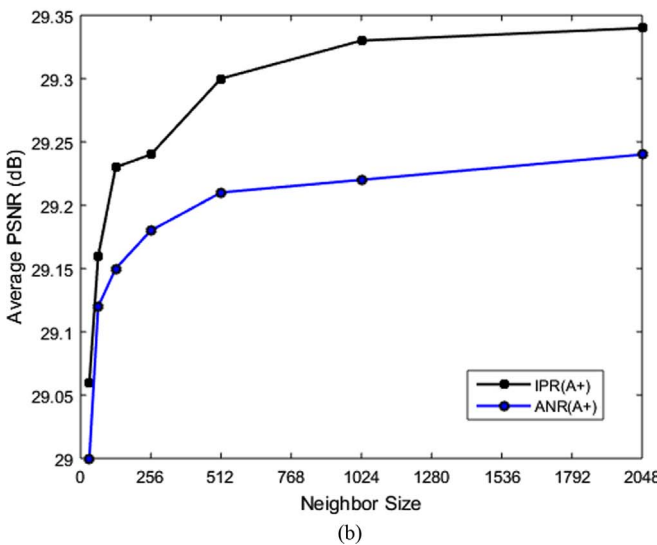
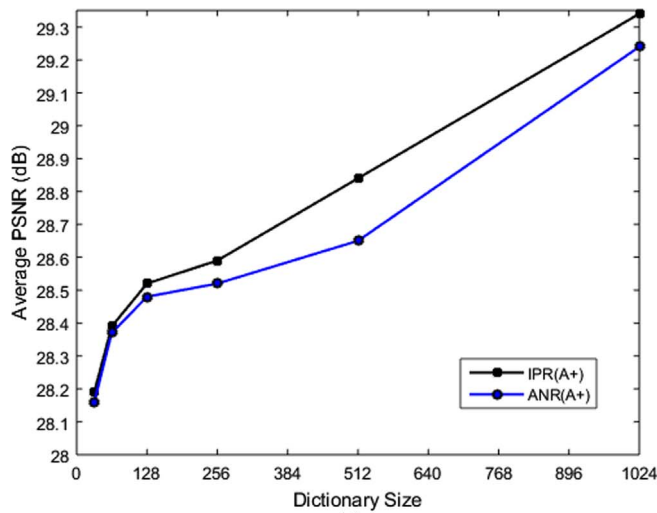
The iterative reconstruction process of the IPR is illustrated in

Fig. 2, which shows that the IPR iteratively refines HF details by adding the reconstructed HR residuals. Note that the LR patch is mean-value-removed in practice, and thus the final reconstructed HR patch should be added to the bicubic upsampling patch to generate the output. It also should be noted that the same set of dictionary atoms are utilized in different iterations. That is because these atoms are mean-value-removed and normalized, they are actually the bases of the residual space. Hence, this set of atoms can also be used to reconstruct the residual components.

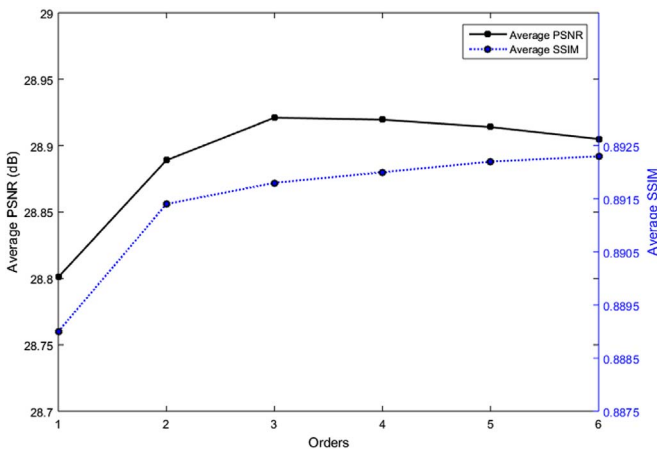
Fig. 3 shows the different upsampling results with the ANR and the proposed IPR for 4X magnification. The test conditions and parameters of the ANR and the IPR are set as the same. It can be found that most of the HF details are missed during the 4X upsampling, and the proposed IPR reproduces relatively finer texture details than the traditional ANR by adding the reconstructed residuals.

(footnote continued)

online supplementary materials: <https://yzhaocv.weebly.com/research.html>



**Fig. 4.** Average PSNR values of ANR(A+) and IPR(A+) methods on ‘Set14’ (3× magnification), (a) with different dictionary sizes (Neighbor size is set as 2048), (b) with different neighbor set sizes (Dictionary size is fixed as 1024).



**Fig. 5.** Average PSNR and SSIM values of IPR (3×) results with different iteration orders on ‘Set14’.

### 3.2. Implementation details

#### 3.2.1. Features

The first- and second-order derivatives of local patch are often used

as the features of LR patch in many example-based methods [26–28,32–36]. Zeyde et al. [32] applied PCA to reduce the dimension of the feature. In this paper, we adopt the same features of LR patch as in [32,33].

#### 3.2.2. Dictionaries and neighborhoods

We utilize the same way to learn the dictionary and to select the neighborhood as in [33,36]. We use the same training set of images proposed by Yang et al. [26], which is also used in many other methods. For the dictionary learning, we adopted the same approach as Zeyde et al. [32], and more details of dictionary learning process can be found in [32]. The traditional ANR and the proposed IPR search the nearest dictionary atom for each LR input and utilize the neighborhood of the nearest atom to pre-compute the projection matrix. The neighborhood of each atom consists of its  $K$  nearest neighbors. Instead of using inner product to search the nearest atom of LR input, Euclidean distance is utilized to measure the distance between different neighbor atoms.

Fig. 4 illustrates the upsampling results in term of different dictionary sizes and different neighborhood sizes respectively. As shown in Fig. 4(a), increasing the dictionary size can obtain higher PSNR values. Similar finding can be found in Fig. 4(b), enlarging the neighbor set is also useful to refine the reconstructed results.

#### 3.2.3. Selection of iterations

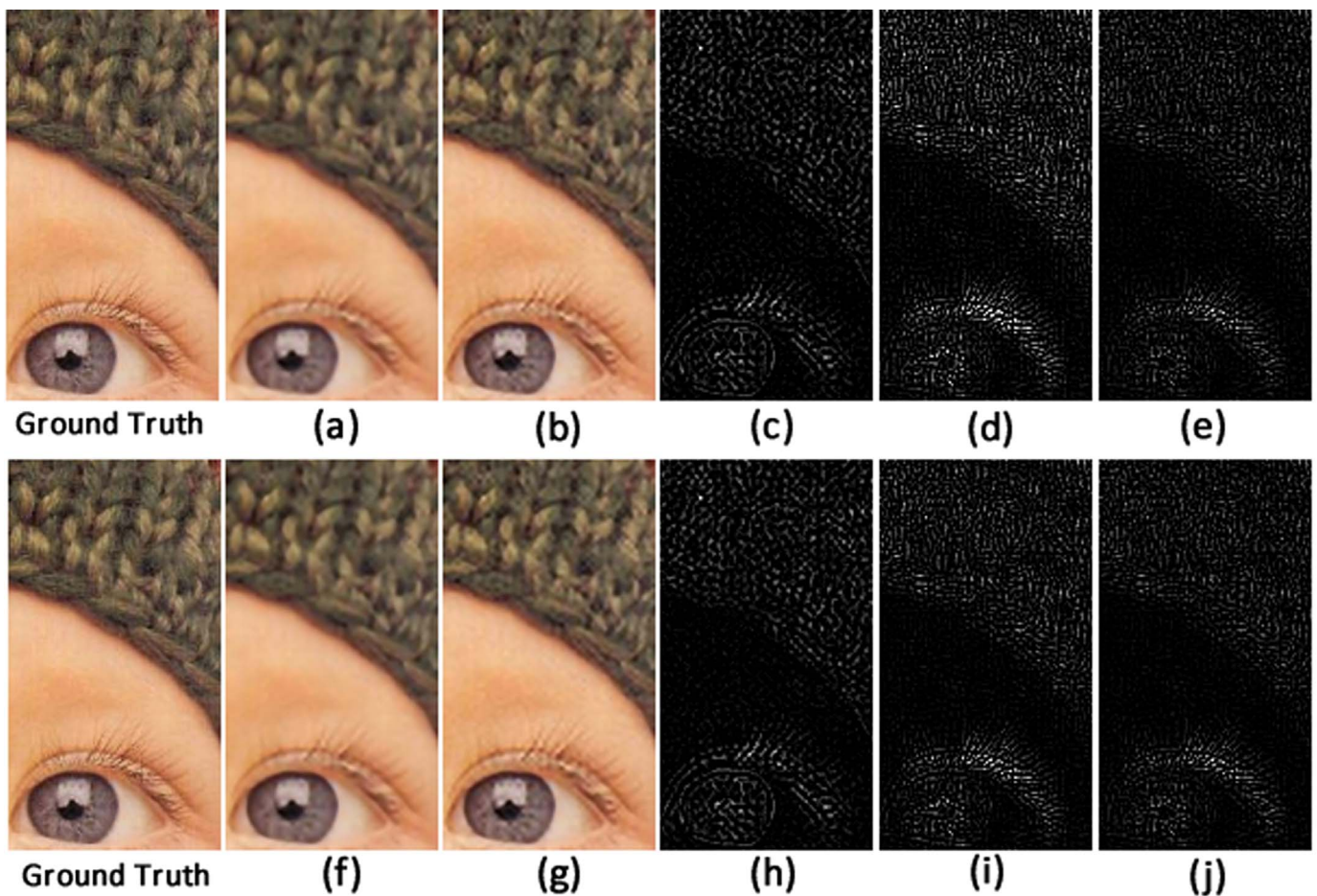
In the proposed IPR, multiple residual components are iteratively used to reconstruct HF details. In the following, we analyze the selection of iterations in terms of time complexity and the objective performance.

The computation time in the proposed IPR method depends on the image size and the number of iterations. Given a LR input feature of size  $n$ , the pre-computed projection matrix of size  $m \times n$ , and the LR dictionary of size  $l \times n$ . The complexity of searching the nearest dictionary atom in each iteration is  $O(ln)$ , and the complexity of multiplying the projection matrix with residual in each iteration is  $O(mn)$ . Hence, the complexity of reconstructing the HR residual in each iteration is  $O(ln+mn)$ . If total  $k$  iterations are computed, the total complexity of reconstructing a HR patch is  $O((k+1)(ln+mn))$ . Therefore, the time cost of the proposed IPR method increases linearly with the number of iterations.

On the other side, the upsampling performance may not always increase along with more iterations. Firstly, the magnitude of the residual component decreases monotonically. The additional HF information thus becomes very small when the number of iterations is large. Secondly, the estimation of the LR patch or residual in each iteration is an inexact approximation. Because the LR patches are estimated by means of fixed atoms, the estimation errors are always exist. Although the proposed IPR tends to reduce the estimation errors by using the iterative process, the errors still cannot be totally avoided. Furthermore, some noises may be accumulated after multiple approximations. Fig. 5 shows the average PSNR and SSIM values with different iterations on image set ‘Set14’. We can see that the PSNR values increase within first three iterations. As the number of iterations unceasingly increases, the PSNR results change very little. The SSIM values are also promoted obviously at the second iterations and then slowly increase along with the augment of iterations. Therefore, the number of iterations are suggested to be set as 2 or 3 by considering both of the performance and the time cost. In this paper, we merely use IPR with 2 iterations in our experiment.

## 4. Experiments

In this section, we compare the proposed IPR with some state-of-the-art methods.



**Fig. 6.** Upsampled results of “baby” image (2×) with the ANR and the proposed IPR, (a) result with the ANR [33], (b) result with the IPR, (c) the residual between the IPR and ANR results, (d) the residual between the ground truth and ANR result, (e) the residual between the ground truth and IPR result, (f) result with the ANR(A+) [36], (g) result with the IPR(A+), (h) the residual between the IPR(A+) and ANR(A+) results, (i) the residual between the ground truth and ANR(A+) result, (j) the residual between the ground truth and IPR(A+) result.

#### 4.1. Testing image sets

We test the proposed method on four testing image sets downloaded from the web and Berkeley image dataset [43], i.e., ‘Set 5’, ‘Set 14’, ‘B100’, and Zhang’s set. ‘Set 5’ [37], ‘Set 14’ [32] and Zhang’s set [49] contain 5, 14, and 9 commonly used images respectively for upsampling evaluation. ‘B100’ [33] consists of 100 testing images selected from the Berkeley image dataset. For color images, they are firstly converted from RGB to YUV. The proposed method and other methods in the comparison are then applied only on Y (intensity) component, and bicubic interpolation is used for U and V components. In our experiments, the input LRIs are obtained by downsampling the original HRIs with bicubic interpolation and the LRIs are then upsampled to their original size with different methods in the comparison respectively. The upsampling factors in our experiment are set as 2, 3, and 4.

#### 4.2. Compared methods and parameters

In this paper, our method is compared with many typical and state-of-the-art methods, such as an interpolation-based methods of the ICBI [11], a reconstruction-based method of Wang’s gradient constrained method [17], and several example-based methods of the Chang’s neighbor embedding method (NE+LLE) [25], the ScSR [26], Zeyde’s method [32], the ASDS [28], Zhang’s dual-dictionary method [49], the ANR [33] and A+ [36]. Note that the A+ is an improvement of the traditional ANR. The only difference between ANR and A+ is that the ANR selects the neighborhood of each atom from the learned dic-

tionary, while the A+ selects larger neighborhood of atom from the original training samples. Hence the proposed IPR can be used directly to improve the A+ in the same way. In this paper, we treat the A+ as a special mode of the ANR rather than a different method. In the experiments, the A+ is denoted by ANR(A+), and the improvement of A+ with the proposed IPR method is denoted by IPR(A+).

In this paper, the dictionary size is 1024 as in [32,33]. For the ANR and IPR, 128 nearest neighbor atoms in the dictionary are selected as the neighborhood of each atom. For the ANR(A+) and IPR(A+), the neighborhood of each atom is composed of 2048 nearest neighbors which are selected from 500,000 training samples. The small number  $\epsilon$  in Eq. (17) is set to  $10^{-5}$ . In this paper, all experiments are performed on an Intel Core i5-3317U laptop PC under MATLAB environment.

#### 4.3. Experimental results

Fig. 6 compares the upsampled results with the ANR and the proposed IPR. Firstly, by comparing the woolen texture in Fig. 6(a), (b), (f), and (g), the proposed IPR and IPR(A+) can recover finer textural details than the original ANR and ANR(A+). Secondly, Fig. 6(d) and (i) illustrate the residuals between ANR methods and the original HRI. The missing HF details, which lead to the blurring texture in ANR methods, mainly locate around tiny edges or textural edges (such as eyelash and wool). Thirdly, the distribution of residual between ANR and IPR results is consistent with the residual between ground truth and ANR results. The added details of IPR method also locate in the texture area, and thus refine the textural details. Overall, the added HF details may not produce very high PSNR promotion

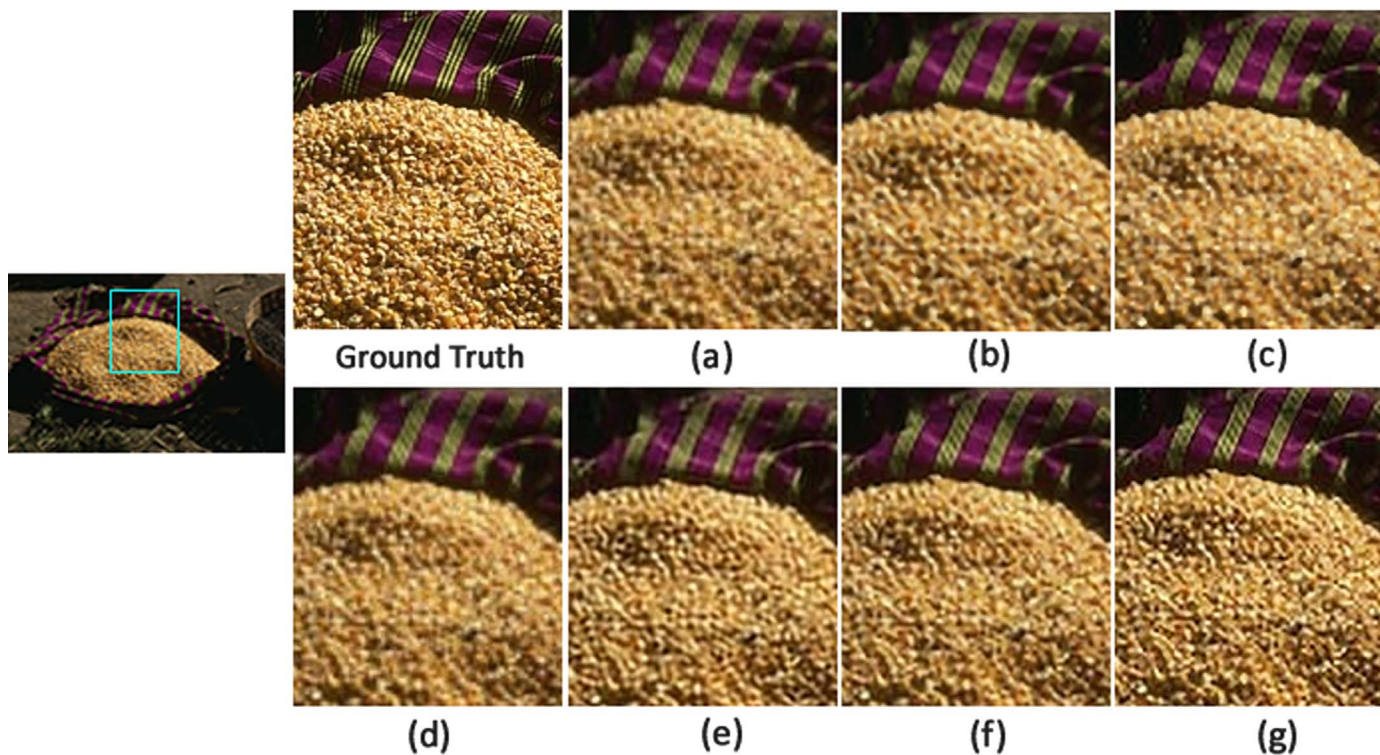


Fig. 7. Upsampled results of “58060” image with different methods (2×), (a) with bicubic interpolation, (b) with the ICBI [11], (c) with Wang’s method [17], (d) with the ScSR [26], (e) with the ASDS [28], (f) with the ANR(A+) [36], (g) with the IPR(A+).

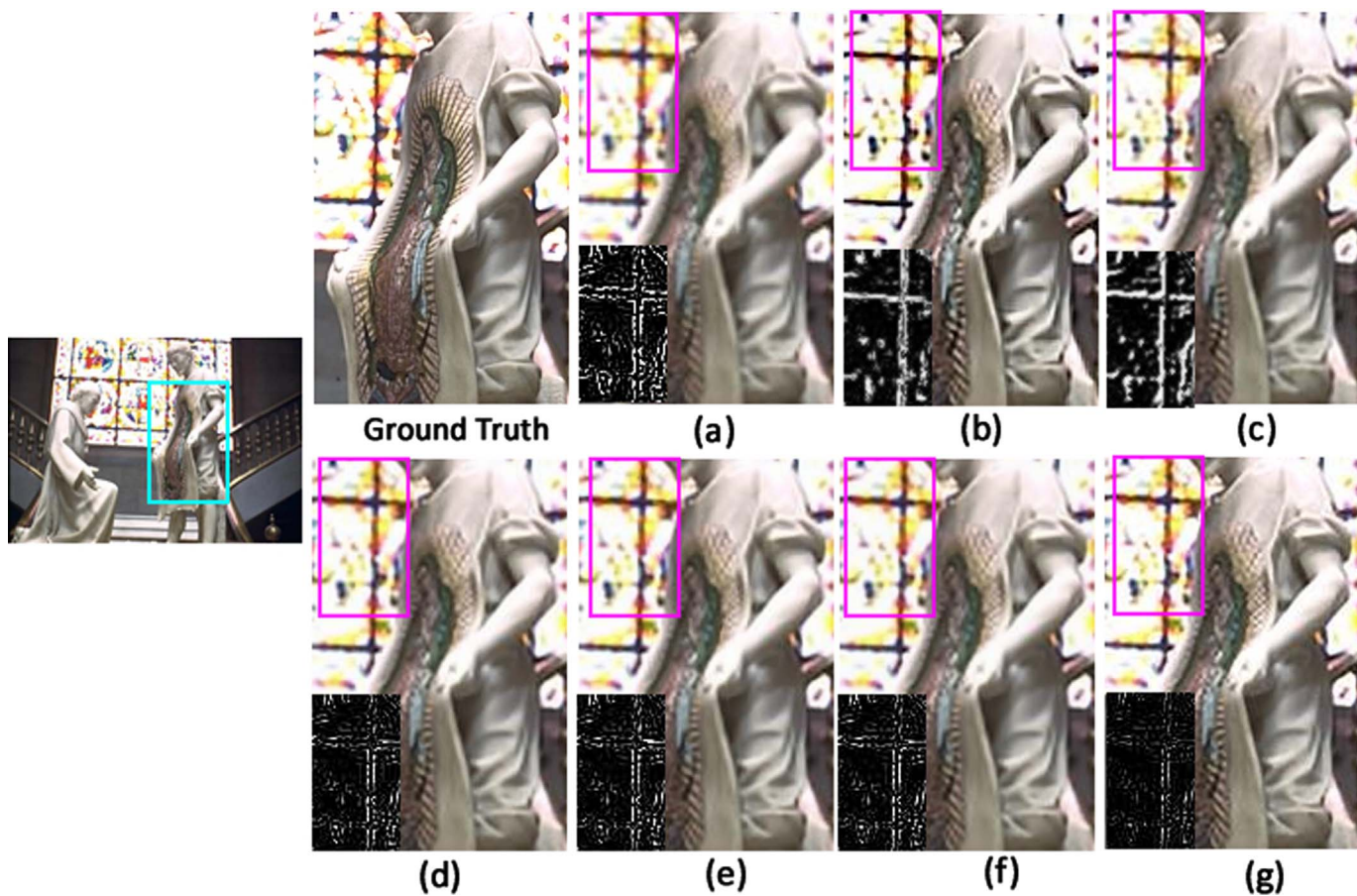
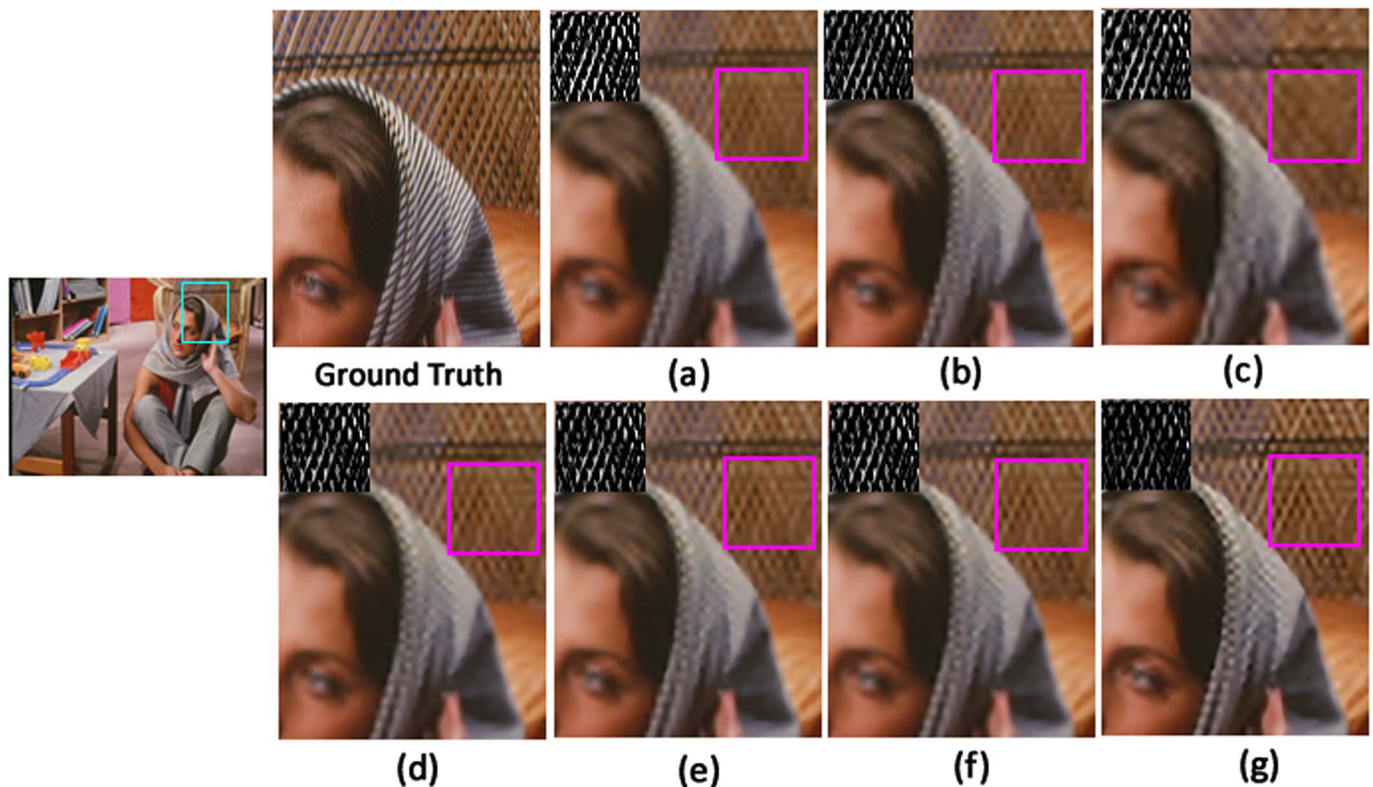


Fig. 8. Upsampled results of “24077” image with different methods (2×), (a) with bicubic interpolation, (b) with the ScSR [26], (c) with the ASDS [28], (d) with Chang’s method [25], (e) with Zeyde’s method [32], (f) with the ANR(A+) [36], (g) with the IPR(A+). The residual between each result and ground truth is shown for the selected area which is marked in the red square.



**Fig. 9.** Upsampled results of “barbara” image with different methods (3×), (a) with bicubic interpolation, (b) with the ScSR [26], (c) with the ASDS [28], (d) with Chang’s method [25], (e) with Zeyde’s method [32], (f) with the ANR(A+) [36], (g) with the IPR(A+). The residual between each result and ground truth is shown for the selected area which is marked in the red square.

because the magnitude of residual is often quite small. However, they can obviously refine the texture since the HF residual can enhance the tiny edges in texture.

Fig. 7 illustrates upsampling results with different methods for 2X magnification. By comparing the close-ups of edges and texture, we can get the following observations. First, the bicubic interpolation produces blurry edge and texture. Although the interpolation-based method ICBI can enhance simple lines and edges, the interpolated texture is still blurry. Second, Wang’s reconstruction-based method can recover sharp edges, but the reconstructed result is sensitive to the parameters. Third, the ANR(A+) and the proposed IPR(A+) reproduce sharper edges than the ScSR and the ASDS. Last, by comparing the details of the grain mass, the IPR(A+) produces more natural texture than the ANR(A+) and other methods.

Fig. 8 compares the proposed IPR with some state-of-the-art example-based methods. We can get some findings. First, all these example-based methods can recover clear edges, and the ANR(A+) and the IPR(A+) reproduce slightly clearer edges than other methods. Second, by comparing the texture on the cloth and window, the IPR(A+) also reconstructs finer texture details than the ANR(A+) and other methods. Last, the residuals between selected area and the ground truth are shown at left bottom of each image. The brighter pixel on the residual map denotes larger difference to original HRI. It is obviously that the proposed IPR(A+) loses the least HF components.

Fig. 9 shows 3× magnification results of ‘barbara’ image with different example-based methods. Comparing the close-ups of the bamboo chair and the cover chief, the IPR(A+) reproduces finer and clearer textural details than other methods. Similar findings can be got from Fig. 10, reconstructed results of these example-based methods become blurry when the upsampling factor increases to 4. The proposed IPR(A+) still recovers the finest texture details and produces the least differences to the original HRIs.

Table 1 lists the image reconstruction performance (average PSNR,

average SSIM, and running time) of various methods on image set ‘Set5’. By comparing the PSNR and SSIM values of the ANR/ANR(A+) and the IPR/IPR(A+), it can be found that the proposed iterative process always performs better than traditional ANR with different upsampling factors. Meanwhile, the proposed IPR(A+) achieves the highest PSNR and SSIM values among these methods. By comparing the running time of these methods, the ANR is the fastest method, and the proposed method is the second fastest method with comparable time cost with the ANR.

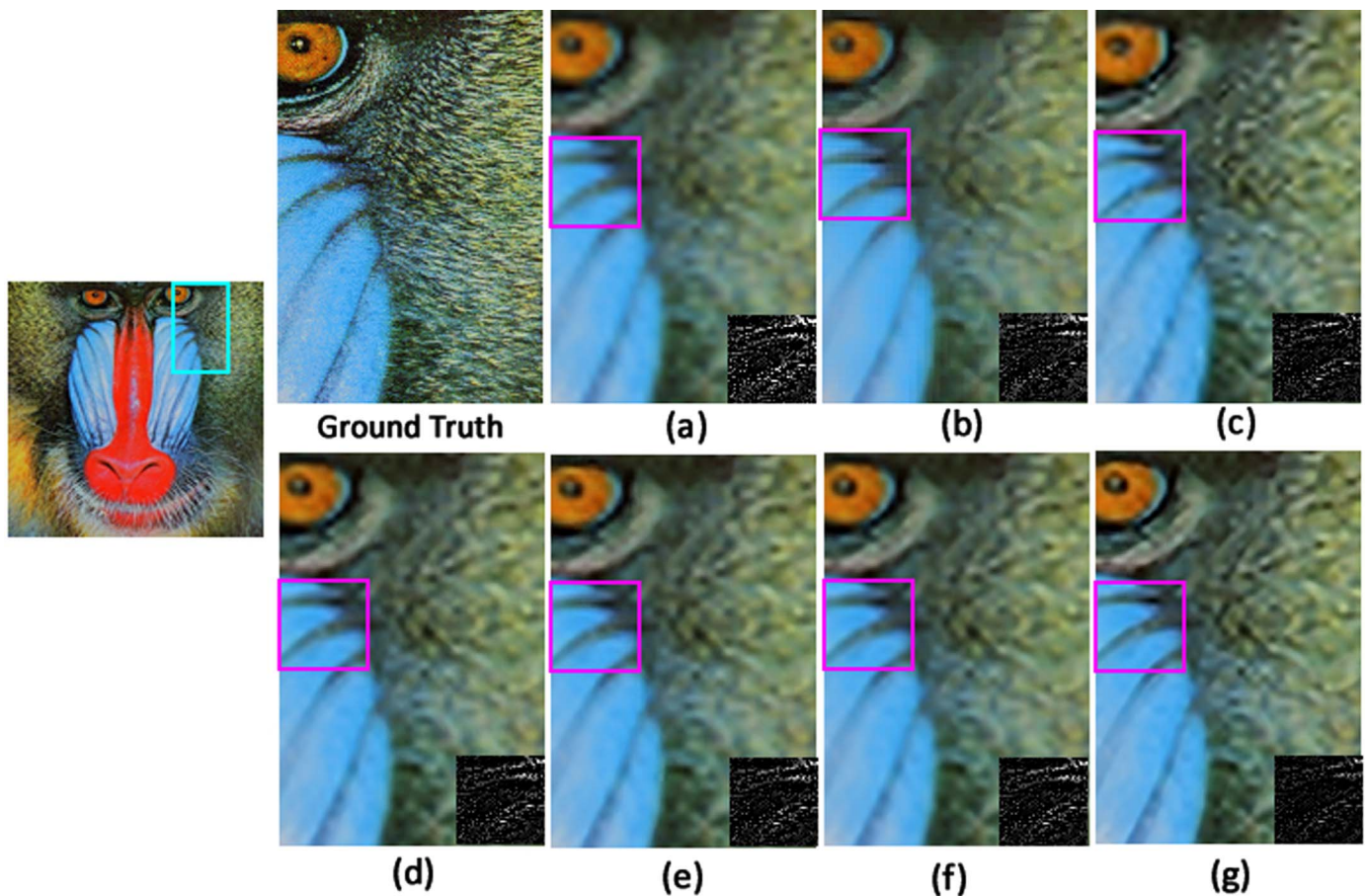
The experimental results of various methods on image sets of ‘Set14’ and ‘B100’ are listed in Table 2 and Table 3, respectively. Similar observations can be got from these two tables. First, the proposed IPR and IPR(A+) can get better objective performance than the ANR and ANR(A+), respectively. Second, the IPR(A+) always outperforms other state-of-the-art example-based methods. Third, the IPR is slightly slower than the ANR due to the additional reconstruction of the residual components, but it is still faster than other example-based methods.

#### 4.4. Comparison with Zhang’s dual-dictionary method

Firstly, we analyze the similarity and the difference between the proposed IPR and Zhang’s method [49]. Both of these methods are focused on the missing HF information which caused by single dictionary-based reconstruction. Hence, both of them tend to refine the HF details through two-fold or multi-fold reconstruction.

However, the two methods are focused on different residual components. In [49], Zhang’s method is focused on the residual between reconstructed result and the ground truth. They thus directly apply the same dictionary-based method after the conventional reconstruction. In the proposed IPR, we focus on the residual during the estimation of input LR patch. An iterative process is then proposed to produce a better estimation of LR input by means of dictionary atoms.





**Fig. 10.** Upsampled results of “baboon” image with different methods (4×), (a) with bicubic interpolation, (b) with the ScSR [26], (c) with the ASDS [28], (d) with Chang’s method [25], (e) with Zeyde’s method [32], (f) with the ANR(A+) [36], (g) with the IPR(A+). The residual between each result and ground truth is shown for the selected area which is marked in the red square.

**Table 1**

Average PSNR (dB), SSIM, and running time (second) of different methods on image set ‘Set5’.

‘Set5’	2×			3×			4×		
	PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
Bicubic	33.66	0.9383	–	30.40	0.8804	–	28.44	0.8250	–
ICBI	32.42	0.9291	44.2	29.13	0.8624	67.2	26.93	0.7995	82.4
Wang’s	35.36	0.9530	7.6	31.14	0.8933	9.3	29.03	0.8488	11.2
NE+LLE	35.57	0.9543	8.4	31.73	0.8836	4.6	29.64	0.8510	3.2
ScSR	35.38	0.9564	226.2	31.23	0.9068	214.3	29.43	0.8551	232.4
ASDS	34.85	0.9544	344.5	31.02	0.9003	331.5	29.54	0.8497	339.2
Zeyde	35.64	0.9559	84.2	31.79	0.9032	43.8	29.69	0.8533	33.1
ANR	35.82	0.9568	0.9	31.84	0.9072	0.7	29.68	0.8558	0.6
IPR	35.93	0.9574	1.4	32.01	0.9077	0.9	29.78	0.8564	0.8
ANR(A+)	36.49	0.9601	0.9	32.57	0.9191	0.7	30.29	0.8732	0.6
IPR(A+)	<b>36.58</b>	<b>0.9615</b>	1.4	<b>32.64</b>	<b>0.9201</b>	0.9	<b>30.35</b>	<b>0.8741</b>	0.8

In addition, the learning algorithms applied in these two methods are slightly different, i.e., the linear regression is used in [49], and the ridge regression is utilized in the proposed method.

Secondly, the comparisons (2× magnification) with Zhang’s method are listed in Table 4. The image set and related implementation details are used as in [49]. By comparing the PSNR values we can find that both of these two methods can obtain fine results by recovering HF information. This also demonstrates the effectiveness of adding missing HF residual components. Furthermore, the proposed IPR(A+) still outperforms this state-of-the-art method.

#### 4.5. Extended experiments and future work

Recently, depth-image upsampling has drawn many attentions [50–52]. We thus implement extended experiments for specific depth-image upsampling scenario. Detailed experimental results can be found in the online supplementary materials.

This work is focused on refining the HF components during the upsampling process. However, it is still very difficult to recover the HF details from a LR input, especially when the magnification factor is large. As illustrated in Fig. 3 and Fig. 10, texture area is totally blurred after the 4X upsampling. Hence, we plan to propose a co-upsampling method to further refine the magnified texture in our future work. In this future work, another similar referenced HR image is selected and

**Table 2**  
Average PSNR (dB), SSIM, and running time (second) of different methods on image set 'Set14'.

'Set14'	2×			3×			4×		
	PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
Bicubic	30.36	0.9417	–	27.67	0.8596	–	26.12	0.7857	–
NE+LLE	31.91	0.9587	18.3	28.74	0.8836	12.5	26.95	0.8137	12.1
ScSR	31.21	0.9620	488.3	28.01	0.8882	562.9	26.57	0.8183	593.2
ASDS	31.15	0.9627	880.7	27.91	0.8938	907.0	26.94	0.8190	870.5
Zeyde	31.96	0.9589	247.1	28.80	0.8841	114.6	26.99	0.8159	131.7
ANR	31.95	0.9626	2.3	28.80	0.8890	2.4	27.00	0.8194	2.0
IPR	32.08	0.9639	3.9	28.89	0.8914	3.7	27.08	0.8226	3.1
ANR(A+)	32.41	0.9641	2.3	29.24	0.8940	2.4	27.44	0.8294	2.0
IPR(A+)	<b>32.50</b>	<b>0.9653</b>	3.9	<b>29.32</b>	<b>0.8966</b>	3.7	<b>27.52</b>	<b>0.8325</b>	3.1

**Table 3**  
Average PSNR (dB), SSIM, and running time (second) of different methods on image set 'B100'.

'B100'	2×			3×			4×		
	PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
Bicubic	29.35	0.8334	–	27.17	0.7361	–	25.95	0.6671	–
NE+LLE	30.40	0.8674	13.7	27.84	0.7687	5.7	26.47	0.6937	3.6
ScSR	30.32	0.8709	284.0	27.74	0.7719	244.3	26.33	0.6997	214.2
ASDS	30.19	0.8712	492.9	27.65	0.7735	433.7	26.45	0.7003	388.2
Zeyde	30.40	0.8682	100.9	27.87	0.7693	68.0	26.51	0.6963	33.6
ANR	30.50	0.8706	2.1	27.90	0.7724	0.9	26.52	0.6991	0.7
IPR	30.78	0.8823	3.8	28.02	0.7777	1.6	26.59	0.7037	1.0
ANR(A+)	30.78	0.8762	2.1	28.18	0.7764	0.9	26.76	0.7062	0.7
IPR(A+)	<b>31.04</b>	<b>0.8843</b>	3.8	<b>28.28</b>	<b>0.7854</b>	1.6	<b>26.83</b>	<b>0.7168</b>	1.1

**Table 4**  
PSNR (dB) comparisons with different algorithms (2×) on Zhang's image set.

Images	Bicubic	Zeyde's	Zhang's [49]	ANR(A+)	IPR(A+)
<i>Foliage</i>	31.65	34.73	35.50	35.59	<b>35.67</b>
<i>Mum</i>	31.05	34.65	35.30	35.40	<b>35.45</b>
<i>Monarch</i>	27.78	30.34	30.88	30.96	<b>31.04</b>
<i>Peppers</i>	32.32	34.46	34.78	34.80	<b>34.84</b>
<i>Flower</i>	32.12	34.97	35.54	35.62	<b>35.76</b>
<i>Window</i>	31.19	33.74	34.20	34.27	<b>34.39</b>
<i>Sailboat</i>	30.56	32.36	32.80	33.20	<b>33.39</b>
<i>Splash</i>	36.16	39.07	39.50	39.53	<b>39.59</b>
<i>Lena</i>	32.19	34.66	34.96	35.03	<b>35.11</b>
<b>Average</b>	31.69	34.33	34.83	34.93	<b>35.03</b>

used to reproduce fine textural details.

## 5. Conclusions

In this paper we proposed a high-performance and low-cost image upsampling method, namely iterative projection reconstruction (IPR). The proposed method utilized an iterative process to recover fine

texture details, and adopted the pre-computed projection matrix to accelerate the reconstruction in each iteration. As a result, the proposed IPR can recover fine texture details with low cost. Experimental results on several image sets demonstrated that the proposed method can obtain better subjective and objective performance compared with state-of-the-art methods.

## Acknowledgements

The authors would like to sincerely thank the anonymous reviewers who have given us so many valuable comments and the suggestion of the future work. We also sincerely thank A. Giachetti, L.F. Wang, W.S. Dong, J.C. Yang, R. Zeyde, J. Zhang, and R. Timofte for sharing the source codes of the ICBI, the Wang's, the ASDS, the ScSR, the Zeyde's, the Zhang's, and the ANR methods. This work was partly supported by the grant of National Science Foundation of China 61370115, 61402018, 61305093, 61305006, 61673157, China 863 project of 2015AA015905, Shenzhen Peacock Plan JCYJ20150331100658943 and JCYJ20160506172227337, and Guangdong Province Project 2014B010117007 for funding.

## Appendix A. Convergence of the IPR algorithm

The  $i$ -th order residue is,

$$\mathbf{R}_i = \mathbf{R}_{i-1} - \langle \mathbf{R}_{i-1}, \mathbf{d}_{q_{i-1}} \rangle \mathbf{d}_{q_{i-1}} \quad (\text{S1-1})$$

Since the  $\mathbf{R}_i$  is orthogonal to the atom  $\mathbf{d}_{q_{i-1}}$ , thus,

$$\|\mathbf{R}_{i-1}\|^2 = \|\mathbf{R}_i\|^2 + |\langle \mathbf{R}_{i-1}, \mathbf{d}_{q_{i-1}} \rangle|^2 \quad (\text{S1-2})$$

We get  $\|\mathbf{R}_i\|^2 < \|\mathbf{R}_{i-1}\|^2$ . Since the weight is monotonically decreased, we have:

$$\omega_i \leq \omega_{i-1} \quad (\text{S1-3})$$

Hence, we have  $\omega_i \|\mathbf{R}_i\|^2 < \omega_{i-1} \|\mathbf{R}_{i-1}\|^2$ .

That means the magnitude of residue components reduces when the order of residue increases, therefore the IPR is convergent.

## References

- [1] R. Keys, Cubic convolution interpolation for digital image processing, *IEEE Trans. Acoust. Speech Signal Process* 29 (6) (1981) 1153–1160.
- [2] T.M. Lehmann, C. Gonner, K. Spitzer, Survey: interpolation methods in medical image processing, *IEEE Trans. Med. Imaging*. 18 (11) (1999) 1049–1075.
- [3] S. Dai, M. Han, W. Xu, Y. Wu, Y. Gong, “ofte edge smoothness prior for alpha channel super resolution, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2007, pp. 1–8.
- [4] X. Li, M.T. Orchard, New edge-directed interpolation, *IEEE Trans. Image Process.* 10 (10) (2001) 1521–1527.
- [5] F. Zhou, W. Yang, Q. Liao, Interpolation-based image super-resolution using multisurface fitting, *IEEE Trans. Image Process.* 21 (7) (2012) 3312–3318.
- [6] D. Su, P. Willis, Image interpolation by pixel-level data-dependent triangulation, *Comput. Graph. Forum* 23 (2) (2004).
- [7] Q. Wang, R.K. Ward, A new orientation-adaptive interpolation method., *IEEE Trans. Image Process.* 16 (4) (2007) 889–900.
- [8] C. Zwart, D. Frakes, Segment adaptive gradient angle interpolation, *IEEE Trans. Image Process.* 22 (8) (2013) 2960–2969.
- [9] X. Liu, D. Zhao, R. Xiong, S. Ma, W. Gao, Image interpolation via regularized local linear regression, *IEEE Trans. Image Process* 20 (12) (2011) 3455–3469.
- [10] Q. Wang, R. Ward, J. Zou, Contrast enhancement for enlarged images based on edge sharpening, in: Proceedings of IEEE International Conference on Image Processing Sep. 2005, vol. 2, pp. 1–4.
- [11] A. Giachetti, N. Asuni, Real-time artifact-free image upscaling, *IEEE Trans. Image Process.* 20 (10) (2011) 2760–2768.
- [12] X. Liu, D. Zhao, J. Zhou, W. Gao, H. Sun, Image interpolation via graph-based bayesian label propagation, *IEEE Trans. Image Process* 23 (3) (2014) 1084–1096.
- [13] M. Irani, S. Peleg, Motion analysis for image enhancement: resolution, occlusion, and transparency, *J. Vis. Commun. Image Represent.* 4 (4) (1993) 324–355.
- [14] H. Stark, P. Oskoui, High resolution image recovery from image-plane arrays, using convex projections, *J. Opt. Soc. Am. A* 6 (11) (2002) 997–1013.
- [15] R. Fattal, Image Upsampling via impose edge statistics, *ACM Trans. Graph.* 26 (3) (2007) (Art. ID 95).
- [16] J. Sun, Z. Xu, H.Y. Shum, Image super-resolution using gradient profile prior, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2008, pp. 1–8.
- [17] L. Wang, S. Xiang, G. Meng, et al., Edge-directed single image super-resolution via adaptive gradient magnitude self-interpolation, *IEEE Trans. Circuits Syst. Video Technol.* 23 (8) (2013) 1289–1299.
- [18] H. Xu, G. Zhai, X. Yang, Single image super-resolution with detail enhancement based on local fractal analysis of gradient, *IEEE Trans. Circuits Syst. Video Technol.* 23 (10) (2013) 1740–1754.
- [19] L. Wang, H. Wu, C. Pan, Fast image upsampling via the displacement field, *IEEE Trans. Image Process.* 23 (12) (2014) 5123–5135.
- [20] Q. Shan, Z. Li, J. Jia, et al., Fast image/video upsampling, *ACM Trans. Graph.* 27 (2008) 32–39.
- [21] T. Michaeli, M. Irani, Nonparametric blind super-resolution, in: Proceedings of IEEE Conference on Computer Vision, Dec. 2013, pp. 945–952.
- [22] N. Efrat, D. Glasner, A. Apartsin, et al., Accurate blur models vs. image priors in single image super-resolution, *Proc. IEEE Int. Conf. Comput. Vis.* (2013) 2832–2839.
- [23] W.T. Freeman, E.C. Pasztor, O.T. Carmichael, Learning low-level vision, *Int. J. Comput. Vis.* 40 (1) (2000) 25–47.
- [24] J. Sun, N.N. Zheng, H. Tao, H. Shum, Image hallucination with primal sketch priors, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2003, vol. 2, pp. 729–736.
- [25] H. Chang, D.Y. Yeung, Y. Xiong, Super-resolution through neighbor embedding, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2004, vol. 1, pp. 275–282.
- [26] J. Yang, J. Wright, T.S. Huang, Y. Ma, Image super-resolution via sparse representation, *IEEE Trans. Image Process* 19 (11) (2010) 2861–2873.
- [27] J. Yang, Z. Wang, Z. Lin, X. Shu, T. Huang, Bilevel sparse coding for coupled feature spaces, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2012, pp. 2360–2367.
- [28] W. Dong, D. Zhang, G. Shi, et al., Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization, *IEEE Trans. Image Process* 20 (7) (2011) 1838–1857.
- [29] J.H. Lee, J.O. Kim, J.W. Han, K.S. Choi, J. Ko, Edge-oriented two-step interpolation based on training set, *IEEE Trans. Consum. Electron.* 56 (3) (2010) 1848–1855.
- [30] K.K. Chua, Y.H. Tay, Enhanced image super-resolution technique using convolutional neural network, *Adv. Vis. Inform.* (2013) 157–164.
- [31] W. Dong, L. Zhang, G. Shi, X. Li, Nonlocally centralized sparse representation for image restoration, *IEEE Trans. Image Process.* 22 (4) (2013) 1620–1630.
- [32] R. Zeyde, M. Elad, M. Protter, On single image scale-up using sparse-representations, *Curv. Surf.* (2010) 711–730.
- [33] R. Timofte, V.D. Smet, L.V. Gool, Anchored neighborhood regression for fast example-based super-resolution, in: Proceedings of IEEE International Conference on Computer Visions, Dec. 2013, pp. 1920–1927.
- [34] L. He, H. Qi, R. Zaretzki, Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Jun. 2013, pp. 345–352.
- [35] T. Peleg, M. Elad, A statistical prediction model based on sparse representations for single image super-resolution, *IEEE Trans. Image Process.* 23 (6) (2014) 2569–2581.
- [36] R. Timofte, V.D. Smet, L.V. Gool, A+: adjusted anchored neighborhood regression for fast super-resolution, in: Proceedings of Asian Conference on Computer Visions, 2014, pp. 1–15.
- [37] M. Bevilacqua, A. Roumy, et al., Low-complexity single image super-resolution based on nonnegative neighbor embedding, in: Proceedings of British Machine Vision Conference, 2012, pp.1–10.
- [38] D. Glasner, S. Bagon, M. Irani, Super-resolution from a single image, in: Proceedings of IEEE International Conference on Computer Visions, Sep. 2009, pp. 349–356.
- [39] G. Freedman, R. Fattal, Image and video upscaling from local self-examples, *ACM Trans. Graph.* 30 (2) (2011) 12–23.
- [40] K. Zhang, X. Gao, D. Tao, X. Li, Single image super-resolution with multiscale similarity learning, *IEEE Trans. Neural Netw. Learn. Syst.* 24 (10) (2013) 1648–1659.
- [41] K. Zhang, D. Tao, X. Gao, X. Li, Z. Xiong, Learning multiple linear mappings for efficient single image super-resolution, *IEEE Trans. Image Process* 24 (3) (2015) 846–861.
- [42] K. Zhang, X. Gao, D. Tao, X. Li, Single image super-resolution with non-local means and steering kernel regression, *IEEE Trans. Image Process* 21 (11) (2012) 4544–4556.
- [43] D. Martin, C. Fowlkes, D. Tal, et al., A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proceedings of IEEE International Conference on Computer Visions, 2001, pp. 416–423.
- [44] Y. Li, C. Cai, G. Qiu, K. Lam, Face hallucination based on sparse local-pixel structure, *Pattern Recognit.* 47 (3) (2014) 1261–1270.
- [45] N. Wang, D. Tao, X. Gao, X. Li, J. Li, A comprehensive survey to face hallucination, *Int. J. Comput. Vis.* 106 (1) (2014) 9–30.
- [46] Y.H. Cohen, R. Fattal, D. Lischinski, Image upsampling via texture hallucination, in: Proceedings of IEEE International Conference on Computational Photography, Mar. 2010, pp. 1–8.
- [47] Y. Zhao, R. Wang, W. Wang, W. Gao, High resolution local structure-constrained image upsampling, *IEEE Trans. Image Process.* 24 (11) (2015) 4394–4407.
- [48] Y. Zhao, R. Wang, W. Wang, W. Gao, Multilevel modified finite radon transform network for image upsampling, *IEEE Trans. Circuits Syst. Video Technol.* (2015). <http://dx.doi.org/10.1109/TCSVT.2015.2504731>.
- [49] J. Zhang, C. Zhao, R. Xiong et al. Image super-resolution via dual-dictionary learning and sparse representation, in: Proceedings of IEEE International Symposium on Circuits and Systems, 2012, pp. 1688–1691.
- [50] Q. Wang, S. Li, H. Qin, A. Hao, Super-resolution of multi-observed RGB-D images based on nonlocal regression and total variation, *IEEE Trans. Image Process.* 25 (3) (2016) 1425–1440.
- [51] D. Ferstl, C. Reinbacher, R. Ranftl, M. R  ther, H. Bischof, Image guided depth upsampling using anisotropic total generalized variation, *Proc. IEEE Int. Conf. Comput. Vis.* (2013) 993–1000.
- [52] A.V. Bhavsar, A.N. Rajagopalan, Resolution enhancement in multi-image stereo, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (9) (2010) 1721–1728.
- [53] K. Zhang, D. Tao, X. Gao, X. Li, J. Li, Coarse-to-fine learning for single-image super-resolution, *IEEE Trans. Neural Netw. Learn. Syst.* 99 (2016) 1–14.
- [54] H. Wang, X. Gao, K. Zhang, J. Li, Single-image super-resolution using active-sampling Gaussian process regression, *IEEE Trans. Image Process.* 25 (2) (2016) 935–948.
- [55] C. Dong, C. Loy, K. He, X. Tang, Image super-resolution using deep convolutional networks, *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (2) (2016) 295–307.
- [56] Z. Cui, H. Chang, S. Shan, et al., Deep network cascade for image super-resolution, in: Proceedings of European Conference on Computer Vision, 2014, pp.49–64.



**Yang Zhao** received the B.E. and the Ph.D. degree from department of automation, University of Science and Technology of China, in 2008 and 2013. From 2013 to 2015, he was a Postdoctoral Researcher with the School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, China. He is currently an Associate Professor with the School of Computer and Information, Hefei University of Technology, China. His research interests include pattern recognition and image processing.



**Rong-Gang Wang** received his Ph.D. Degree from the Institute of Computing Technology, Chinese Academy of Sciences. He was a Research staff in Orange (France telecom) Labs from 2006 to 2010. He is currently an Associate Professor with Peking University Shenzhen Graduate School. His research interest is on video coding and processing. He has done many technical contributions to ISO/IEC MPEG and China AVS. He led the MPEG Internet Video Coding (IVC) standard, served as MPEG IVC AHG co-chair since 2012, and served as AVS implementation sub-group co-chair since 2015. He has authored more than 50 papers in international journals and conferences, and held more than 40 patents.



**Wei Jia** received the B.Sc. degree in informatics from Central China Normal University, Wuhan, China, in 1998, the M.Sc. degree in computer science from Hefei University of Technology, Hefei, China, in 2004, and the Ph.D. degree in pattern recognition and intelligence system from University of Science and Technology of China, Hefei, China, in 2008. He has been an assistant and associate professor in Hefei Institutes of Physical Science, Chinese Academy of Science from 2008 to 2015. He is currently an associate professor in School of Computer and Information, Hefei University of Technology. His research interests include computer vision, biometrics, pattern recognition, image processing and machine learning.



**Wen-Min Wang** received the B.E. degree in Computer Application from Harbin University of Science and Technology in 1983, then received the M.E. and the Ph.D. degrees in Computer Architecture from Harbin Institute of Technology, China, respectively in 1986 and 1989. From 1992, he had 18 years of oversea industrial experience in Japan and America. He came back academia from the end of 2009, currently is a professor with the School of Electronic and Computer Engineering of Peking University, China. His current research interests include video analysis, multimedia retrieval, artificial intelligence and machine learning.



**Wen Gao** received the Ph.D. degree in electronics engineering from the University of Tokyo, Japan, in 1991. He is a Professor of computer science with Peking University, China. Before joining Peking University, he was a Professor of computer science with the Harbin Institute of Technology from 1991 to 1995, and a Professor with the Institute of Computing Technology of Chinese Academy of Sciences. He has published extensively including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics. He served on the editorial board for

several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON AUTONOMOUSMENTAL DEVELOPMENT, the *EURASIP Journal of Image Communications*, the *Journal of Visual Communication and Image Representation*. He chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ICME and ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations.