

## Research Article

# Wavelet Domain Multidictionary Learning for Single Image Super-Resolution

Xiaomin Wu,<sup>1</sup> Jiulun Fan,<sup>1</sup> Jian Xu,<sup>1,2</sup> and Yanzi Wang<sup>1</sup>

<sup>1</sup>*School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China*

<sup>2</sup>*Image Processing and Recognition Center, Xi'an Jiaotong University, Xi'an 710049, China*

Correspondence should be addressed to Jian Xu; [xujian\\_paper@126.com](mailto:xujian_paper@126.com)

Received 22 October 2014; Accepted 27 April 2015

Academic Editor: Zabih F. Ghassemlooy

Copyright © 2015 Xiaomin Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Image super-resolution (SR) aims at recovering the high-frequency (HF) details of a high-resolution (HR) image according to the given low-resolution (LR) image and some priors about natural images. Learning the relationship of the LR image and its corresponding HF details to guide the reconstruction of the HR image is needed. In order to alleviate the uncertainty in HF detail prediction, the HR and LR images are usually decomposed into 4 subbands after 1-level discrete wavelet transformation (DWT), including an approximation subband and three detail subbands. From our observation, we found the approximation subbands of the HR image and the corresponding bicubic interpolated image are very similar but the respective detail subbands are different. Therefore, an algorithm to learn 4 coupled principal component analysis (PCA) dictionaries to describe the relationship between the approximation subband and the detail subbands is proposed in this paper. Comparisons with various state-of-the-art methods by experiments showed that our proposed algorithm is superior to some related works.

## 1. Introduction

Image super-resolution (SR) is a technology of generating a high-resolution (HR) image from one or more low-resolution (LR) images. SR has many applications such as criminal investigation and medical diagnosis.

SR approaches can be roughly classified into three categories: interpolation-based methods, reconstruction-based methods, and example-based methods. Interpolation-based methods [1–4], such as bicubic interpolation, are efficient, but they tend to lose many details of the images. Reconstruction-based methods [5–7] usually assume some degraded models and solve the inverse problems to obtain the HR images. These methods perform well when the magnification factor is small, but their performance depends heavily upon a rational prior imposed on the required images. Learning-based methods [8–24] assume that high-frequency (HF) details lost in an LR image can be recovered by learning the relationships between the LR-HR image patch-pairs. This kind of methods shows robust SR capability when a larger

magnification ratio is required. In this paper, we focus on the learning-based method.

Recently, various typical learning-based SR methods have been proposed. Most of these methods realize SR by inferring the mapping relationships between the HR and LR image patch-pairs. For example, Freeman et al. [8] applied a belief propagation algorithm to train the Markov network which was utilized to model the relationships between HR and LR patch pairs. However, this method will introduce some unwanted artifacts. By introducing locally linear embedding (LLE) [25] from manifold learning, Chang et al. [9] proposed a neighbor-embedding (NE) algorithm. It assumes that the two manifolds of the LR image patches and the corresponding HR patches have similar local structures. However, the manifold assumption is not always true for the SR problem because there are one-to-many mappings existing between the LR-HR patches. Inspired by the fact that the image patches can be well-represented as a sparse linear combination of elements from an appropriately chosen overcomplete dictionary, Yang et al. [11] proposed a sparse-coding-based SR

by employing a relationship that the LR-HR image patch-pairs share with the same sparse coefficients with respect to their own dictionaries. However, the dictionary training process is time-consuming. Zeyde et al. [12] introduced some modifications to Yang's [11] algorithm. They adopted a different approach to train the LR-HR dictionary pairs and generated training data sets with no access to an external set of images. Timofte et al. [15] accelerated the SR based on sparse representation by gathering the dictionary atoms together. Nazzal et al. [23, 24] proposed two SR methods by training three or four HR-LR pairs of dictionaries in the wavelet domain, respectively. It is clear that these methods are time-consuming. Yang et al. [16] also proposed a fast regression model for practical SR based on in-place examples by leveraging two fundamental SR methods, learning from an external database and learning from self-examples (a patch in the upper scale image has good matches around its origin location in the lower scale image). This method has low time complexity but will introduce some unwanted artifacts. A Bayesian method constrained by a beta process prior is applied to learn the overcomplete dictionaries by He et al. [19]. Chen and Qi [20] proposed an SR method based on the low-rank matrix recovery and neighbor embedding. Xu et al. [21] proposed a new coupled K-SVD dictionary training method for image super-resolution. It is a more efficient method of dictionary training and largely improves the performance of K-SVD. Xu et al. [22] also proposed another useful method for image super-resolution by mid-frequency sparse representation and total variation regularization.

The most important task of image SR is to recover the HF details of an HR image according to the LR image or its prior assumptions. However, the existing methods do not well exploit the information provided by input image for this task. We use 1-level discrete wavelet transform (DWT) to decompose a given image into 4 subbands, including an approximation subband and three detail subbands called horizontal, vertical, and diagonal detail subbands.

There are two reasons to do this decomposition:

- (1) Most of the sparse representation based algorithms are working on local patches. DWT can well extract the local spatial and frequency information of the images.
- (2) The relationship between these subbands can be learned from the natural images. The relationship can be utilized in the reconstruction.

Figure 1(a) shows the 1-level DWT result of the ground HR image. It is clear that the approximation subbands of the original HR image and the bicubic interpolated image (Figure 1(b)) of the corresponding LR image after 1-level DWT are very similar but the detail subbands are different. That means the approximation subband of the bicubic interpolated image can be served as the approximation subband of the resultant image. The detail subbands can be reconstructed by the relationship recorded by the coupled dictionaries. Therefore, we introduce the sparse coding method to train the relationships of the subbands.

Though the universal dictionaries learned from example image patches by using algorithms such as K-SVD [26] are

better adapted to local image structures, the K-SVD gets reliable learning only when the number of the data samples is much larger than the dimension of the dictionary. Therefore, the process of dictionary-learning by K-SVD is time-consuming. In this paper, we train 4 principal component analysis (PCA) dictionaries which can well represent 4 subbands, respectively. Every dictionary is composed of a set of compact subdictionaries which is learned from high-quality training patches. The training patches are clustered into many clusters. Since each cluster consists of many patches with similar patterns, a compact subdictionary can be learned for each cluster. Thus, the PCA subdictionary-learning method does not need a large number of training patches. The atoms obtained by PCA are orthogonal to each other. These relevant atoms will not only decrease the computational cost in sparse coding but also will improve the representation accuracy. The initial reconstruction image of our method and its corresponding subbands can be seen in Figure 1(c). From this figure, we can see that many textures in the detail subbands are reconstructed by sparse representation technology.

The rest of this paper is organized as follows. Section 2 briefly reviews the related work of our algorithm. Section 3 describes the details of the proposed wavelet domain SR method. The experimental parameter settings and the results are given in Section 4. The last section concludes our paper.

## 2. Related Works

*2.1. 2D Discrete Wavelet Transform (DWT).* When digital images are to be viewed or proposed at multiple resolutions, the DWT is the mathematical tool of choice. As an efficient, highly intuitive framework for the representation and storage of multiresolution images, the DWT provides powerful insight into an image's spatial and frequency characteristics. The Fourier transform, however, reveals only an image's frequency attributes.

1-level DWT of a gray-scale image can be presented as follows:

$$\mathbf{w}(m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \mathbf{I}(x, y) \varphi(x, y), \quad (1)$$

where  $(M, N)$  is the size of image  $\mathbf{I}(x, y)$ .  $\varphi(x, y)$  is the scale function.  $\mathbf{w}(m, n)$  is the approximation subimage of  $\mathbf{I}(x, y)$  and  $(m, n)$  is the position of each pixel. Consider

$$\mathbf{w}^i(m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \mathbf{I}(x, y) \psi^i(x, y) \quad (2)$$

$$i \in \{H, V, D\},$$

where  $\psi^i(x, y)$  and  $i \in \{H, V, D\}$  are the 2-dimensional wavelets and the superscript  $i$  is a set of directional variables named horizontal, vertical, and diagonal directions.  $\mathbf{w}^i(m, n)$  represents detail subimages corresponding to different directional variables.

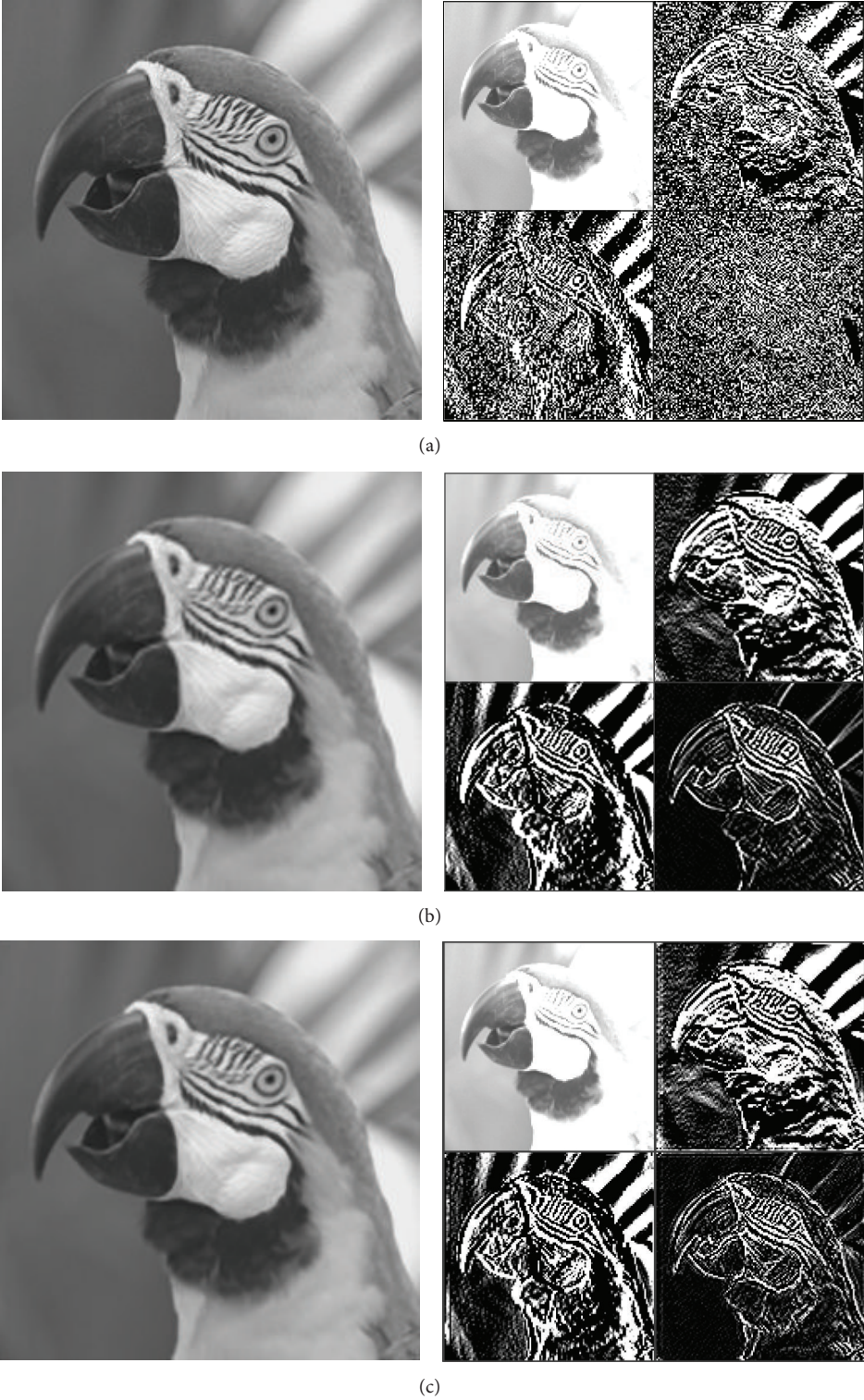


FIGURE 1: The images and their corresponding decomposition images after 1-level DWT (the subimages contain an approximation subband and three detail subbands). (a) The ground HR parrot image and its decomposition images. (b) The bicubic interpolated image and its DWT subband images. (c) The HR image reconstructed by our method and its DWT subband images.



**2.2. Sparse Representation Based SR Method.** The observed LR image patch is a blurred and downsampled version of the HR image patch  $\mathbf{x}$ :

$$\mathbf{y} = F(\mathbf{x}), \quad (3)$$

where  $F$  is a function which can describe the relationship of the LR-HR image patch pairs.

The goal of SR is to recover  $\mathbf{x}$  from  $\mathbf{y}$ . Because SR remains extremely ill-posed, a sparse prior is used in order to regularize it in this paper. By jointly training two dictionaries for the LR and HR image patches, we enforce the similarity of the sparse representations between the LR and HR image patch pairs with respect to their own dictionaries. Thus, for each input LR patch  $\mathbf{y}_j$ , we can find a sparse representation with respect to LR dictionary  $\mathbf{D}_l$ . The corresponding HR patch  $\mathbf{x}_j$  can be reconstructed by the HR dictionary  $\mathbf{D}_h$  and the sparse representation  $\alpha_j$ . The joint sparse coding model can be formulated as [27]:

$$\begin{aligned} \min_{\mathbf{D}_l, \mathbf{D}_h, \alpha_j} \quad & \sum_j \|\mathbf{D}_l \alpha_j - \mathbf{y}_j\|_2^2 + \|\mathbf{D}_h \alpha_j - \mathbf{x}_j\|_2^2 + \lambda \|\alpha_j\|_1 \\ \text{s.t.} \quad & \|\mathbf{D}_l(:, p)\|_2 \leq 1, \\ & \|\mathbf{D}_h(:, p)\|_2 \leq 1, \end{aligned} \quad (4)$$

where  $\alpha_j$  represents the  $j$ th sparse coefficient vector and  $\mathbf{D}_l(:, p)$  is the  $p$ th column of  $\mathbf{D}_l$ .  $\|\cdot\|_1$  is the  $l_1$ -norm of a vector.  $\|\cdot\|_2$  is the  $l_2$ -norm of a vector and  $\lambda$  is a Lagrange parameter balancing the sparsity penalty and representation fidelity.

### 3. Proposed Algorithm

In this section, we describe the specific processes of our proposed algorithm; and the flowchart of our proposed algorithm is shown in Figure 2.

**3.1. Training Four Subband Dictionaries.** According to the aforementioned content, the goal of the training phase is to learn 4 dictionaries corresponding to the subimages in the wavelet domain. Because the human visual system is more sensitive to the luminance component than the chrominance components, we use the YCbCr color space for color images and only perform SR reconstruction in the luminance component.

Firstly, for an HR luminance image  $\mathbf{X}_s(x, y)$ , we decompose it into four subimages named approximation subband  $\mathbf{w}_s$  and three detail subbands  $\mathbf{w}_s^i$ ,  $i \in \{H, V, D\}$  by 1-level DWT as follows:

$$\begin{aligned} \mathbf{w}_s(m, n) &= \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \mathbf{X}_s(x, y) \varphi(x, y), \\ \mathbf{w}_s^i(m, n) &= \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \mathbf{X}_s(x, y) \psi^i(x, y) \end{aligned} \quad (5)$$

$i \in \{H, V, D\}.$

Secondly, we extract the gradient features of  $\mathbf{w}_s$  and divide the feature images into patches and reshape them into the columns. For each detail subband, we directly partition it into patches. Now we obtain four patch matrices,  $\mathbf{W}_s$ ,  $\mathbf{W}_s^H$ ,  $\mathbf{W}_s^V$ , and  $\mathbf{W}_s^D$ , by stacking the corresponding columns.

One important issue of sparsity-based SR is the selection of the dictionary. In this paper, we train 4 principal component analysis (PCA) dictionaries which can represent 4 subbands, respectively. Comparing with the dictionaries generated by the famous algorithm K-SVD [26], the PCA dictionary-learning method does not need a large number of training patches, and the atoms of the dictionary are orthogonal to each other. The small dictionary size saves the memory and time costs. In addition, since a given patch can be better represented by the adaptively selected subdictionary, the whole image can be more accurately reconstructed than using a universal dictionary. The same idea has been validated by Dong et al. [13]. So we classify the training patches into  $K$  clusters and learn a principal component analysis (PCA) subdictionary for each cluster. These  $K$  PCA subdictionaries construct a large overcomplete dictionary to characterize all the possible local structures of natural images.

That means for the  $k$ th cluster of the training patches  $\mathbf{W}_{s_k}$ ,  $\mathbf{W}_{s_k}^i$  ( $k = 1, \dots, K$ ),  $i \in \{H, V, D\}$ , the joint sparse coding model which is used to train the PCA subdictionaries is as follows:

$$\begin{aligned} \min_{\mathbf{D}_{s_k}^i, \mathbf{A}_{s_k}^i} \quad & \|\mathbf{W}_{s_k} - \mathbf{D}_{s_k} \mathbf{A}_{s_k}\|_F^2 + \|\mathbf{W}_{s_k}^i - \mathbf{D}_{s_k}^i \mathbf{A}_{s_k}^i\|_F^2 \\ & + \sum_j \|\alpha_j^k\|_0, \\ \text{s.t.} \quad & \|\mathbf{D}_{s_k}(:, p)\|_2 \leq 1, \\ & \|\mathbf{D}_{s_k}^i(:, p)\|_2 \leq 1, \end{aligned} \quad (6)$$

where  $\mathbf{D}_{s_k}$  is the  $k$ th PCA subdictionaries corresponding to approximation subimage and the atoms of this dictionary are the  $P_k$  principal components of the patch matrix  $\mathbf{W}_{s_k}$ .  $\mathbf{A}_{s_k} = \{\alpha_j^k\}_{j=1, \dots, P_k}$  is the sparse coefficient matrix by calculating  $\mathbf{A}_{s_k} = \mathbf{D}_{s_k}^T \mathbf{W}_{s_k} \cdot \|\cdot\|_0$  is the number of the nonzero elements of a vector. The detail subdictionaries  $\mathbf{D}_{s_k}^i$ ,  $i \in \{H, V, D\}$  corresponding to  $\mathbf{D}_{s_k}$  can be reformulated as follows:

$$\mathbf{D}_{s_k}^i = \arg \min_{\mathbf{D}_{s_k}^i} \|\mathbf{W}_{s_k}^i - \mathbf{D}_{s_k}^i \mathbf{A}_{s_k}^i\|_F^2. \quad (7)$$

The solution to (7) is given by the following pseudoinverse expression:

$$\mathbf{D}_{s_k}^i = \mathbf{W}_{s_k}^i \mathbf{A}_{s_k}^{i+} = \mathbf{W}_{s_k}^i \mathbf{A}_{s_k}^{i T} (\mathbf{A}_{s_k}^i \mathbf{A}_{s_k}^{i T})^{-1}. \quad (8)$$

Thus, we can train  $4K$  subdictionaries corresponding to 4 subimages by (6) and (8).

**3.2. Test Phase.** The test phase attempts to magnify an LR test image  $\mathbf{Y}$ . From Figure 1, we assume that the approximation subimages of the desired HR image and the bicubic

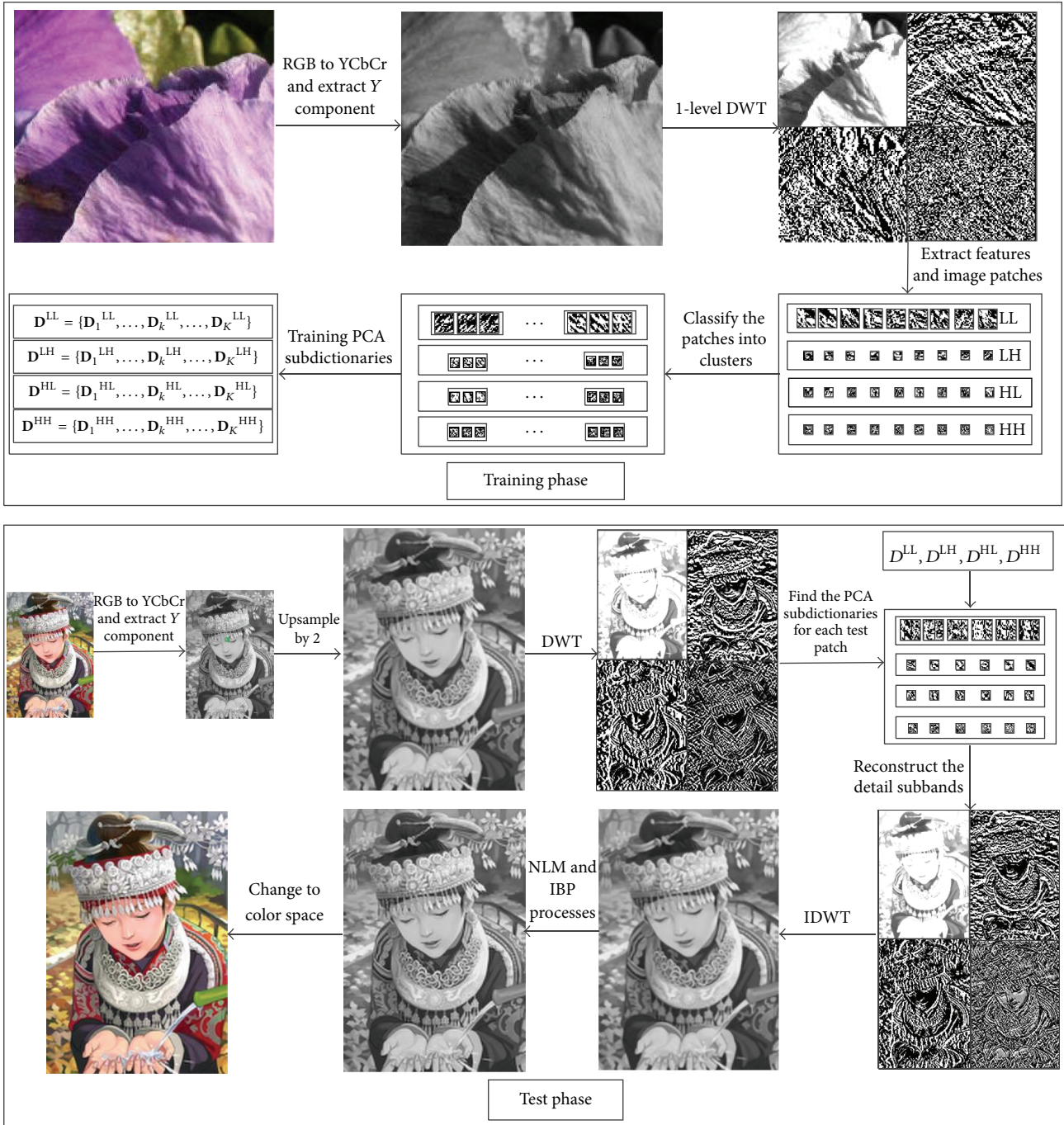


FIGURE 2: The flowchart of our proposed algorithm.

interpolated image are very similar. We magnify the test LR image by a factor of 2 and transform the bicubic interpolated image by 1-level DWT to get the four subimages. Then the approximation subband of it can be simply served as the approximation subband of the required HR image. The goal of this phase is to reconstruct the other three detail subbands from the approximation subband.

Firstly, we need to extract the gradient features of the approximation subimage. We partition the feature images into patches and reshape each of them as column of the matrix  $W_t = \{y_t\}_{t=1}^T$ , where  $T$  is the number of patches.

Secondly, for each patch to be coded, we adaptively select one subdictionary from the trained  $K$  approximation subdictionaries to code it. This actually enforces the coding

coefficients of this patch over the other subdictionaries to be 0, leading to a very sparse representation of the given patch. So the sparse representation  $\alpha_t$  of each patch  $\mathbf{y}_t$  with respect to the learned approximation subdictionary  $\mathbf{D}_{s_k}$  is the solution to the problem:

$$\alpha_t = \mathbf{D}_{s_k}^T \mathbf{y}_t. \quad (9)$$

Thirdly, the detail subband patches corresponding to can be computed by the following equation:

$$\mathbf{y}_t^i = \mathbf{D}_{s_k}^i \alpha_t, \quad i \in \{H, V, D\}. \quad (10)$$

Three subimages can be produced by merging the corresponding patches. For the overlapping regions between those adjacent patches, averaging fusion is applied to obtain the estimated pixels.

In the end, the initial HR image  $\mathbf{X}_0$  will be obtained by inverse DWT of the approximation subband and the received 3 detail subbands.

**3.3. Postprocessing Phase.** The initial output HR image always loses some details in the input LR image and contains some artifacts. In order to make the image more suitable for human visual system, we need to fit the reconstructed HR image to the input LR image. Because the recently introduced non-local redundancies in image are useful for image denoising and restoration [28], we can incorporate the nonlocal self-similarities in our postprocessing phase. Therefore, a global constraint such as Iterative Back-Projection (IBP) [29] and a nonlocal similarity constraint are employed to further improve the reconstruction accuracy as follows [19]:

$$\begin{aligned} \mathbf{X}^* &= \underset{\mathbf{X}}{\text{Arg min}} \quad \|\mathbf{X} - \mathbf{X}_0\|_2^2 \\ \text{s.t.} \quad &\mathbf{Y} = \mathbf{D}\mathbf{B}\mathbf{X}, \\ &\left\| \mathbf{x}_j - \sum_{s=1}^S b^s \bar{\mathbf{x}}_j^s \right\|_2 \leq \xi, \end{aligned} \quad (11)$$

where  $\mathbf{x}_j$  and  $\bar{\mathbf{x}}_j$  are patches in  $\mathbf{X}$  and  $\mathbf{X}_0$ , respectively.  $\bar{\mathbf{x}}_j^s$  ( $s = 1, \dots, S$ ) is the  $s$ th most similar patch to  $\bar{\mathbf{x}}_j$  and  $b^s$  is the nonlocal weight defined in [28].  $\xi$  is the permitted error of the truth patch and the weighted evaluated patches.  $\mathbf{D}$  and  $\mathbf{B}$  are downsampling and blurring operations, respectively.

The summary of the training phase can be found in Algorithm 1.

*Algorithm 1* (the training phase of the wavelet domain super-resolution).

**Input:**

A group of training images;

**Output:**

Trained approximation subdictionaries:

$$\mathbf{D}_s = \{\mathbf{D}_{s_1}, \dots, \mathbf{D}_{s_k}, \dots, \mathbf{D}_{s_K}\};$$

The detail subdictionaries:

$$\mathbf{D}_s^i = \{\mathbf{D}_{s_1}^i, \dots, \mathbf{D}_{s_k}^i, \dots, \mathbf{D}_{s_K}^i\}, \quad i \in \{H, V, D\};$$

Cluster centroid matrix  $\mathbf{C}$ ;

- (1) Transform the training image  $\mathbf{Y}_s$  by 1-level DWT and extract the gradient features of the approximation subband;
- (2) Partition the feature-images and the other three detail subimages into patches in raster-scan order to construct the training data sets  $\mathbf{W}_s, \mathbf{W}_s^H, \mathbf{W}_s^V, \mathbf{W}_s^D$ ;
- (3) Cluster the patches in  $\mathbf{W}_s$  into  $K$  clusters by  $K$ -means, and save the cluster centroid matrix as  $\mathbf{C}$ ;
- (4) Train PCA subdictionaries  $\mathbf{D}_{s_k}$  ( $k = 1, \dots, K$ ) by (6) and the corresponding detail subdictionaries  $\mathbf{D}_{s_k}^i$ ,  $i \in \{H, V, D\}$  by (8).

In addition, the summary of the test phase is listed in Algorithm 2 as follows.

*Algorithm 2* (the test phase of the wavelet domain super-resolution).

**Input:**

Test image  $\mathbf{Y}$ ;

Trained PCA subdictionaries:

$$\mathbf{D}_s = \{\mathbf{D}_{s_1}, \dots, \mathbf{D}_{s_k}, \dots, \mathbf{D}_{s_K}\},$$

$$\mathbf{D}_s^i = \{\mathbf{D}_{s_1}^i, \dots, \mathbf{D}_{s_k}^i, \dots, \mathbf{D}_{s_K}^i\}, \quad i \in \{H, V, D\};$$

Cluster centroid matrix  $\mathbf{C}$ ;

**Output:**

HR image  $\mathbf{X}^*$ ;

- (1) Upscale the test image to obtain the bicubic interpolated image  $\mathbf{Y}_0$ . Transform  $\mathbf{Y}_0$  by 1-level DWT and name the approximation subimage as  $\mathbf{Y}_1$ ;
- (2) Extract the gradient features of  $\mathbf{Y}_1$  to generate the feature images. Partition the feature images into patches in raster-scan order to construct the test data set  $\mathbf{Y}_t$ ;
  - (i) Find the PCA subdictionary  $\mathbf{D}_{s_k}$  for each test patch  $\mathbf{y}_t$  in  $\mathbf{Y}_t$  as follows:
$$\arg \min_k \|\mathbf{y}_t - \mathbf{C}(:, k)\|_2^2, \quad k = 1, \dots, K, \quad (12)$$

where  $\mathbf{C}(:, k)$  is the  $k$ th column of the matrix  $\mathbf{C}$ ;
  - (ii) Calculate the sparse representation  $\alpha_t$  for each  $\mathbf{y}_t$  by (9) and the corresponding patches of the detail subimages by (10);
- (3) Generate the detail subimages by sequentially merging all patches of each detail subimage, respectively;
- (4) Get the initial HR image  $\mathbf{X}_0$  by inverse DWT of the  $\mathbf{Y}_1$  and the received three detail subimages;
- (5) Perform the nonlocal self-similarity and IBP on  $\mathbf{X}_0$  to obtain the final HR image  $\mathbf{X}^*$  by (11).



TABLE 1: The PSNR and SSIM values of the proposed method with different wavelet-filters (for each wavelet, the left is PSNR values and the right is SSIM values).

Img.	haar		db4		sym4		bior6.8	
Baboon	<b>25.60</b>	<b>0.76</b>	25.50	0.76	25.55	0.76	25.54	0.76
Comic	<b>27.88</b>	<b>0.90</b>	27.52	0.89	27.78	0.90	27.77	0.90
Flower	<b>32.44</b>	<b>0.93</b>	32.02	0.93	32.32	0.93	32.31	0.93
Girl	35.64	<b>0.89</b>	35.63	0.89	35.64	0.89	<b>35.65</b>	0.89
Lena	<b>36.43</b>	<b>0.93</b>	36.17	0.93	36.35	0.93	36.33	0.93
Man	<b>30.55</b>	<b>0.88</b>	30.23	0.88	30.43	0.88	30.41	0.88
Pepper	<b>36.78</b>	<b>0.92</b>	36.27	0.92	36.63	0.92	36.61	0.92
Zebra	<b>33.50</b>	<b>0.94</b>	32.92	0.94	33.36	0.94	33.36	0.94
Parrot	<b>33.79</b>	<b>0.96</b>	33.08	0.95	33.49	0.95	33.38	0.95
Parthenon	<b>29.21</b>	<b>0.85</b>	28.82	0.84	29.06	0.84	29.02	0.84
Raccoon	<b>32.50</b>	<b>0.90</b>	32.32	0.90	32.43	0.90	32.39	0.90
Avg.	<b>32.21</b>	<b>0.90</b>	31.86	0.89	32.09	0.89	32.07	0.89

#### 4. Experimental Results and Analysis

In this section we analysis the experimental settings and results.

*4.1. Experimental Settings.* In our experiments, we downloaded the software package for [11] as our HR training images and 69 images were used to prepare our training database; and we randomly selected some benchmark images which were often used in others' works to prepare our test database. For all of our training and test images, we use the YCbCr color space for color images and only perform our method in the luminance component. The chrominance components can be upscaled to the desired size with the bicubic interpolation method. We select 150000 textured patches with fine details as our training examples to avoid uninformative patches affecting the efficiency of learning.

The settings of parameters always affect the performance of our method. To get the best results, we discuss some important parameters involved in this method.

Firstly, we need to know which wavelet in the wavelet family is the best for our proposed method; the Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measurement (SSIM) [30] values of each test image when the number of clusters is 370 and the patch size set 3 with 2 overlapping pixels are shown in Table 1. From Table 1, we can draw a conclusion that the best wavelet of our algorithm is *haar* wavelet. On one hand, shift variance of the wavelet transform is considered as the major drawback for the applications in texture analysis [31]. Except for *haar* filter, the strict shift invariance of the wavelet transform is not achievable. On the other hand, the phase linearity (symmetry) of the decomposition is also an important condition in classification. Nonsymmetric filters will lead to the degradation of the image and lower accuracy in classification due to the nonlinear phase. Therefore, we choose *haar* wavelet in our experiments.

Secondly, we cluster the training patches extracted from a set of example images into  $K$  clusters by  $K$ -means algorithm;

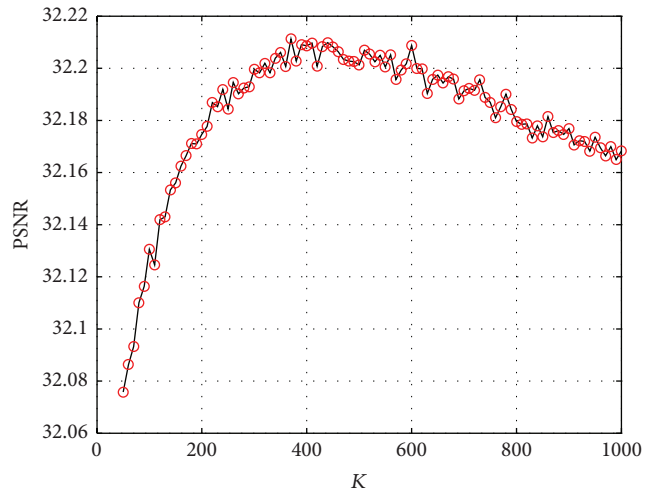


FIGURE 3: The average PSNR values of all test images versus different numbers of clusters.

and for each test patch, we compute the distances of it with the centroids to determine its cluster and corresponding PCA subdictionary. To verify the number of clusters  $K$ , we set it from 50 to 1000 with a step of 10 on all test images when the patch size is 3 with 2 overlapping pixels. From Figure 3, we can see that, when the number of clusters is smaller than 300, the PSNR values are small. The results indicate that, when the number of the cluster is small, the texture in a cluster is too complex for a compact representation produced by PCA; and, when  $K$  is bigger than 610, a small amount of the data samples is considered as a group, which makes the proposed method work like NE. The average PSNR value reaches its peak value when  $K = 370$ . Therefore, we use 370 in our following experiments.

Finally, our experiments all work on image patches. We partition the image into patches with some overlapping pixels

TABLE 2: The comparisons of the proposed algorithm with other methods by a magnification factor of 2 (for each method, the values are PSNR and SSIM from left to right).

Img.	Bicubic		Dong [13]		Zeyde [12]		Yang [16]		ANR [15]		Our	
Baboon	24.87	0.70	25.04	0.71	25.47	0.75	25.53	0.76	25.54	0.76	<b>25.60</b>	<b>0.76</b>
Comic	26.01	0.85	27.05	0.87	27.65	0.90	27.41	0.89	27.77	0.90	<b>27.88</b>	<b>0.90</b>
Flower	30.39	0.89	31.41	0.90	32.27	0.92	31.83	0.92	<b>32.51</b>	0.93	32.44	<b>0.93</b>
Girl	34.76	0.87	34.45	0.84	35.50	0.88	35.61	0.89	35.60	0.89	<b>35.64</b>	<b>0.89</b>
Lena	34.71	0.91	35.33	0.91	36.21	0.93	36.06	0.93	36.32	0.93	<b>36.43</b>	<b>0.93</b>
Man	29.25	0.85	29.78	0.85	30.44	0.88	30.14	0.88	30.47	0.88	<b>30.55</b>	<b>0.88</b>
Pepper	35.09	0.91	35.87	0.90	36.59	0.92	36.04	0.92	36.39	0.92	<b>36.78</b>	<b>0.92</b>
Zebra	30.65	0.91	32.16	0.92	33.21	0.94	32.6	0.94	33.07	0.94	<b>33.50</b>	<b>0.94</b>
Parrots	31.28	0.94	32.18	0.93	33.27	0.95	32.90	0.95	33.59	0.95	<b>33.79</b>	<b>0.96</b>
Parthenon	28.08	0.81	28.38	0.80	29.15	0.84	28.82	0.84	29.06	0.84	<b>29.21</b>	<b>0.85</b>
Raccoon	30.95	0.85	31.00	0.85	32.21	0.89	32.33	0.90	32.45	0.90	<b>32.50</b>	<b>0.90</b>
Avg.	30.55	0.86	31.15	0.86	32.00	0.89	31.75	0.89	32.07	0.89	<b>32.21</b>	<b>0.90</b>

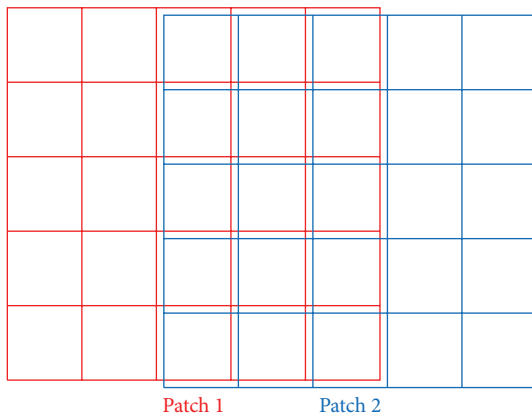


FIGURE 4: One block represents a pixel, and the patch with red solid curve is overlapped partitioned by the adjacent patch with blue solid curve. The overlapping pixels of each two adjacent patches are  $2 * 5 = 10$ .

in raster-scan order. An example of partitioning an image into patches with the patch size of 5 and overlapping pixels of 2 is shown in Figure 4. The overlapping pixels between the adjacent patches may ensure the reconstructed image with consistency and avoid blocking effect. Therefore, to measure the size of image patch and overlapping pixels, we set the patch size from 2 to 5 with different overlapped pixels on all the test images while fixing the number of clusters to 370; and the average PSNR values corresponding to different patch sizes with different overlapping pixels are shown in Figure 5. Many related works have indicated the reputation property of the local patches [10, 17]. The smaller the patch is, the better reputation is [10]. That means it is easier to find similar patches in the training data set if the patch size is smaller. Therefore, smaller patch obtains higher PSNR values. Larger overlap provides smoother results, which meets the smooth prior of the natural images. So the size of image patch is set to 3 with 2 pixels overlapped between the adjacent patches.

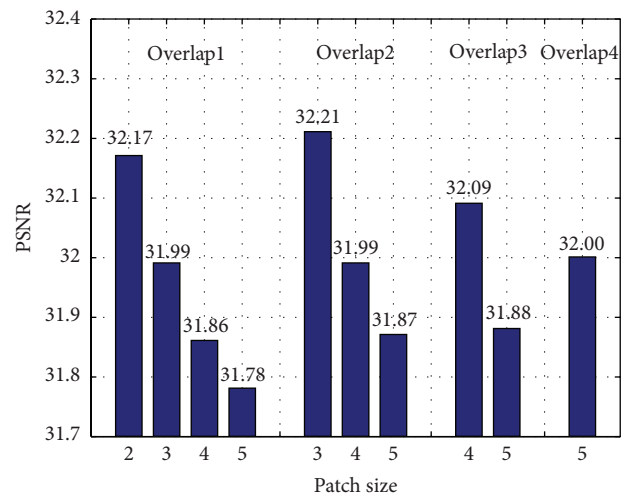


FIGURE 5: The average PSNR values with respect to different patch sizes with different overlap pixels.

## 4.2. Experimental Comparisons

**4.2.1. Quantitative Comparisons.** We compare the performance of the proposed algorithm with other algorithms [12, 13, 15, 16] by a magnification factor of 2. For the fair comparison, the algorithms of [13, 16] use the same training examples as ours. We use PSNR and SSIM [30] to evaluate the performance of reconstruction. The results of all the algorithms are shown in Table 2. As shown, our proposed method gets the highest SSIM values, and most of the PSNR values are higher than other methods.

**4.2.2. Qualitative Comparisons.** In order to further describe the effectiveness of our algorithm, the visual comparison results with other algorithms by a factor of 2 are shown in Figures 6 and 7. It is obvious that the proposed algorithm does not introduce undesired artifacts like ringing but produces rich texture details and sharp edges.





FIGURE 6: The comparison of the proposed algorithm and four other methods on *lena* image (2x), and the local magnification in green rectangle is shown in the upper right corner in each example. (a) Bicubic. (b) Zeyde et al. [12]. (c) Yang et al. [16]. (d) ANR [15]. (e) Our proposed. (f) Original.

**4.3. Computation Complexity.** We analyze the computational complexity and CPU execution time of our test phase.

Supposed that there are  $S$  patches in the test set and the dimension of the patch is  $d^2$ , the dimension of the gradient feature of test patch is  $4d^2$ . For each sample, we find the nearest cluster by computing the Euclidian distance of this example with the centroids which is obtained in the training process. Suppose the size of the nearest cluster is  $k_i$ , the computation complexity of finding the cluster is  $O(4k_i d^2)$ . Then we need to compute the sparse representation vector of this example by (9). Suppose the size of the PCA subdictionary with respect to the nearest cluster is  $k_i^p$ , the computation complexity of computing the sparse representation is  $O(4k_i^p d^2)$ . In addition, the three detail patches can be calculated by (10), and the computation complexity of these steps is  $O(3(k_i^p)^2 d)$ . In conclusion, the computation complexity for all  $S$  test patches is  $O(S(4k_i d^2 + 4k_i^p d^2 + 3(k_i^p)^2 d))$ .

For better understanding, the CPU execution time comparisons with some state-of-the-art methods of test phase are shown in Figure 8. The configurations of our computer are as follows: AMD FX (tm)-8120 Eight-Core Processor, 8G RAM, and 64-bit operating system. It is clear that our method is

faster than SCSR (Sparse Coding Super-resolution) [11] and slightly slower than Zeyde et al.'s [12].

## 5. Conclusion and Future Works

This paper proposes a novel algorithm about single image SR based on training PCA dictionaries in wavelet domain. We decompose the image into 4 subbands, which obtains good performance by alleviating the uncertainty in HF details prediction. In addition, the PCA dictionary provides low time complexity in both the training and testing phases. In the future, we will upscale each subband with different method according to its own characteristics. The dictionary group obtained by the proposed method is trained with the training sets about one magnification factor, so it is only suitable for one magnification factor. Therefore, how to train dictionaries to fit different magnification factors is still a problem.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

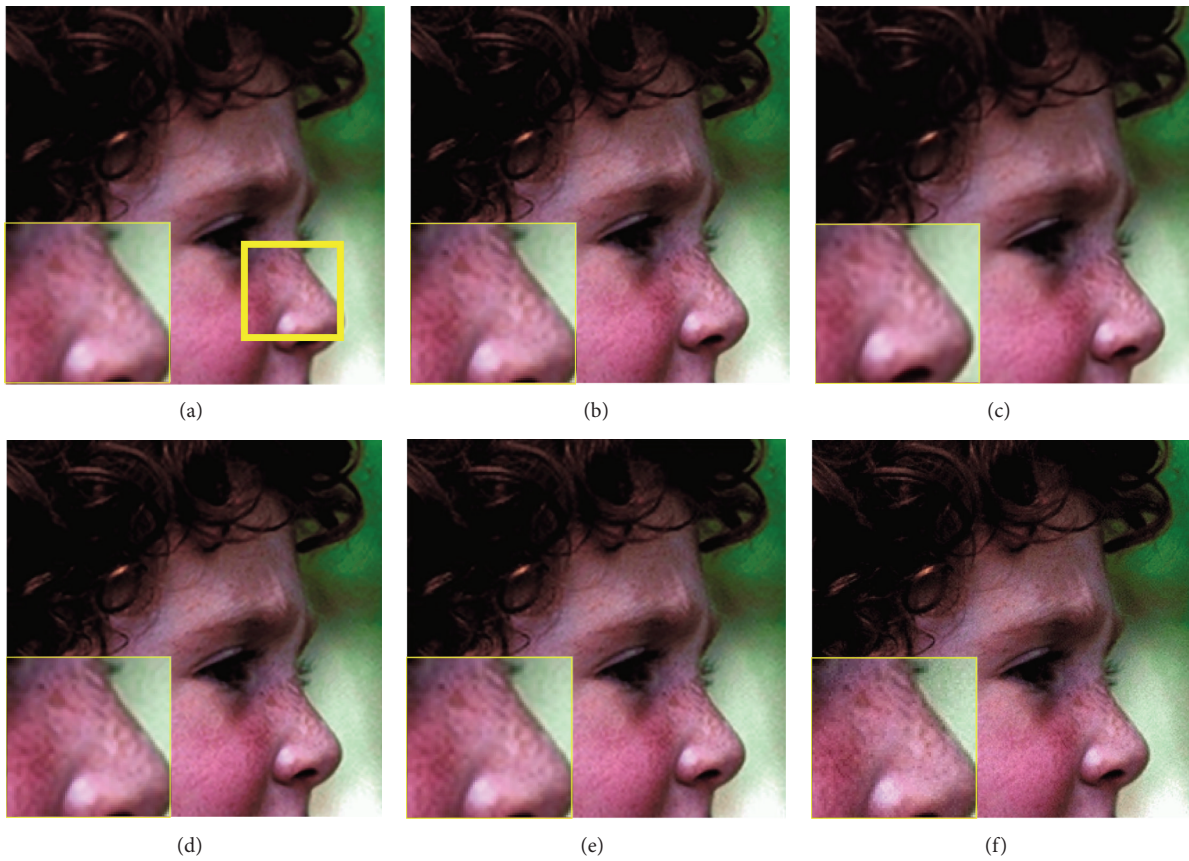


FIGURE 7: The comparison of the proposed algorithm and four other methods on *girl* image (2x), and the local magnification in yellow rectangle is shown in the lower left corner in each example. (a) Bicubic. (b) Zeyde et al. [12]. (c) Yang et al. [16]. (d) ANR [15]. (e) Our proposed. (f) Original.

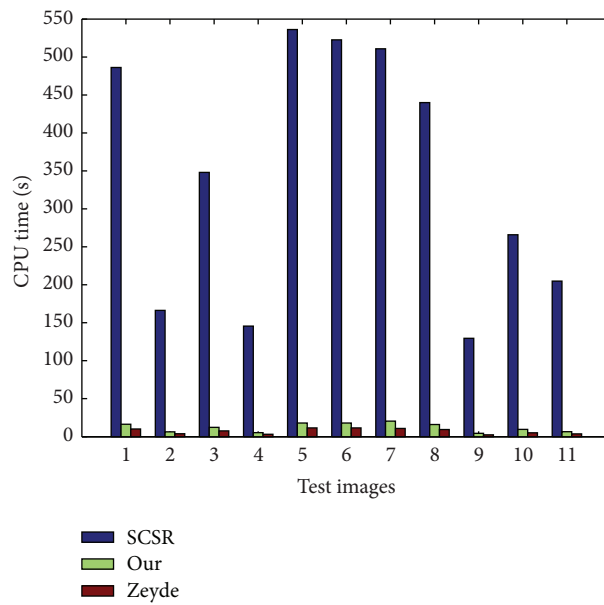


FIGURE 8: The CPU time comparison of our method with [11, 12].

## Acknowledgments

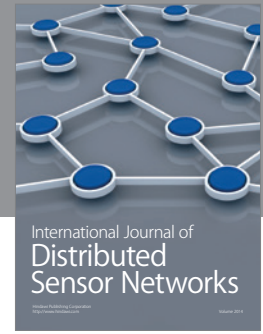
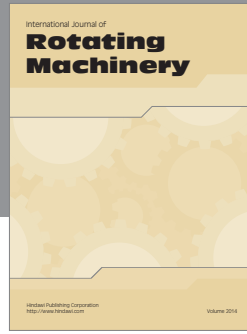
This work was supported by the National Science Foundation of China (Grant nos. 61340040, 61202183, and 61102095) and Young Teachers Foundation of Xi'an University of Posts and Telecommunications.

## References

- [1] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Transactions on Image Processing*, vol. 17, no. 6, pp. 887–896, 2008.
- [2] O. A. Omer and T. Tanaka, "Image superresolution based on locally adaptive mixed-norm," *Journal of Electrical and Computer Engineering*, vol. 2010, Article ID 435194, 4 pages, 2010.
- [3] S. Mallat and G. Yu, "Super-resolution with sparse mixing estimators," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2889–2900, 2010.
- [4] H. Chavez-Roman and V. Ponomaryov, "Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 10, pp. 1777–1781, 2014.
- [5] J. Sun, Z. Xu, and H.-Y. Shum, "Image super-resolution using gradient profile prior," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, AK, USA, June 2008.
- [6] Y.-W. Tai, S. Liu, M. S. Brown, and S. Lin, "Super resolution using edge prior and single image detail synthesis," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2400–2407, June 2010.
- [7] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with non-local means and steering kernel regression," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4544–4556, 2012.
- [8] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 56–65, 2002.
- [9] H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, pp. 1275–1282, July 2004.
- [10] D. Glasner, S. Bagon, and M. Irani, "Super-resolution from a single image," in *Proceedings of the IEEE 12th International Conference on Computer Vision (ICCV '09)*, pp. 349–356, October 2009.
- [11] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [12] R. Zeyde, M. Elad, and M. Protter, "On single image scaleup using sparse-representations," in *Proceedings of the 7th International Conference on Curves and Surfaces*, pp. 711–730, Avignon, France, June 2010.
- [13] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, 2011.
- [14] X. Gao, K. Zhang, D. Tao, and X. Li, "Image super-resolution with sparse neighbor embedding," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3194–3205, 2012.
- [15] R. Timofte, V. De, and L. V. Gool, "Anchored neighborhood regression for fast example-based super-resolution," in *Proceedings of the 14th IEEE International Conference on Computer Vision (ICCV '13)*, pp. 1920–1927, December 2013.
- [16] J. Yang, Z. Lin, and S. Cohen, "Fast image super-resolution based on in-place example regression," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 1059–1066, June 2013.
- [17] K. Zhang, X. Gao, D. Tao, and X. Li, "Single image super-resolution with multiscale similarity learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 10, pp. 1648–1659, 2013.
- [18] S. Ram and J. J. Rodriguez, "Single image super-resolution using dictionary-based local regression," in *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI '14)*, pp. 121–124, San Diego, Calif, USA, April 2014.
- [19] L. He, H. Qi, and R. Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 345–352, Portland, Ore, USA, June 2013.
- [20] X. Chen and C. Qi, "Low-rank neighbor embedding for single image super-resolution," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 79–82, 2014.
- [21] J. Xu, C. Qi, and Z. Chang, "Coupled K-SVD dictionary training for super-resolution," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '14)*, pp. 3910–3914, Paris, France, October 2014.
- [22] J. Xu, Z. Chang, J. Fan, X. Zhao, X. Wu, and Y. Wang, "Image superresolution by midfrequency sparse representation and total variation regularization," *Journal of Electronic Imaging*, vol. 24, no. 1, Article ID 013039, 2015.
- [23] M. Nazzal and H. Ozkaramanli, "Improved single image super-resolution using sparsity and structured dictionary learning in wavelet domain," in *Proceedings of the 21st Signal Processing and Communications Applications Conference*, pp. 1–4, Haspolat, Turkey, April 2013.
- [24] M. Nazzal and H. Ozkaramanli, "Wavelet domain dictionary learning-based single image superresolution," *Signal, Image and Video Processing*, pp. 1–11, 2014.
- [25] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [26] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [27] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled dictionary training for image super-resolution," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3467–3478, 2012.
- [28] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 2, pp. 60–65, June 2005.



- [29] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, no. 3, pp. 231–239, 1991.
- [30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [31] A. Mojsilovic, M. V. Popovic, and D. M. Rackov, "On the selection of an optimal wavelet basis for texture characterization," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2043–2050, 2000.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

