An Adaptive Super-Resolution Image Reconstruction Algorithm Based on Image Blocks

Yong Yin[1*], Qianqian Ruan[1], Shitong Liu[1] and Xiaoli Chen[1]

[1]College of Communication Engineering, Chongqing University, Chongqing 400044, China
*yongyin@cqu.edu.cn

Abstract

Recently there have been significant advances in image super-resolution, we embark from an method to learn an over-complete dictionary with K-SVD algorithm, adopted a two-layer dictionary learning method because of the reason of restriction of over-complete dictionary's size and the intrinsic sparsity of the algorithm, and introduced collaborative representation to replace the sparse representation in reconstruction phase of two-layer dictionary learning method. And combined with an image reconstruction algorithm based on singular value decomposition, proposed an adaptive super-resolution image reconstruction algorithm based on image blocks. The proposed algorithm take advantage of high arithmetic speed of adaptive two-layer dictionary learning method and high image reconstruction precision of image reconstruction with singular value decomposition by which adaptive choose different method to reconstruct image blocks. The experimental results show that the proposed algorithm can greatly decrease the image reconstruction time and barely reduce the reconstruction precision.

***Keywords**: super-resolution image reconstruction, sparse representation, two-layer dictionary learning, adaptive*

1. Introduction

Single-image super-resolution(SR) is a popular branch of image reconstruction that focuses on the enhancement of image resolution to generate a visually pleasing high-resolution output image from a low-resolution input image. SR is an ill-posed problem because each low resolution(LR) pixel has to be mapped onto many high resolution(HR) pixels, depending on the desired up-sampling factor.

Example-based super-resolution methods have been an active research topic to recover high- frequency details with the help of a database consisting of co-occurrence examples from a training set of HR and LR image patches [1-4]. Recent achievements in sparse representation suggest that the linear relationships among high-dimension signals can be accurately recovered from their low-dimension projections [5]. Instead of working directly with the patch pairs sampled from high and low-resolution images, based on the assumption that for a given low-resolution patch, the sparse representation vector in the over-complete dictionary trained by low-resolution images is the same as the one of its corresponding high-resolution patch in the over-complete dictionary trained by high-resolution images, learn a compact representation for these patch pairs to capture the co-occurrence prior to improve the speed and the robustness significantly, achieving state-of-art performance [7]. However, due to the restriction of over-complete dictionary's size and the intrinsic sparsity of the algorithm, one limitation in recovering high-frequency details should be noticed. As the gap between the frequency spectrum of the corresponding HR image and that of the initial interpolation is showed that

---

* Corresponding Author

learning-based algorithm usually can't work well. To address above problem, a two-layer algorithm is proposed, in which details of high-frequency are estimated by a progressive way [8]. The scheme can be seen as a coarse-to-fine recovering process and better results can be expected. However, the method increased computational complexity as the same time. In this paper, we adopted collaborative representation to replace the sparse representation in reconstruction phase, thus we can compute a projection matrix with the dictionary trained offline, then project the low-resolution image patch to high-resolution image patch, which reduced computation greatly. An image reconstruction is introduced to make use of rank deficiency, which exploited the corresponding relation between the sparse and low rank after singular value decomposition to image [11]. The singular value decomposition of image matrix uses the property of low rank for image reconstruction to obtain a pleasing reconstructed result. However, it has high computational complexity.

To make use of adaptive two-layer dictionary method's high executive speed and image reconstruction based on singular value decomposition method's high reconstruction precision, we proposed an adaptive of image reconstruction algorithm based on image blocks, which can choose the algorithm adaptively based on image block feature of variance to improve the speed of image super-resolution reconstruction.

In the following section we will described the two-layer dictionary learning method simply, which introduced collaborative representation to replace the sparse representation in reconstruction phase, and the image reconstruction algorithm based on singular value decomposition in Section 2. In Section 3, an adaptive image reconstruction algorithm is introduced and is explained in detail. All the experimental results in Section 4 and the conclusions are drawn in Section 5.

2. Dictionary-Based Super-Resolution Image Reconstruction

2.1. Two-Layer Dictionary Super-Resolution Image Reconstruction

The scheme also consists of a dictionary learning phase that trains dual dictionaries, namely main dictionary and residual dictionary and an image reconstruction phase, performing the image super-resolution on the input image using the trained model from the previous phase, as described in Figure 1.
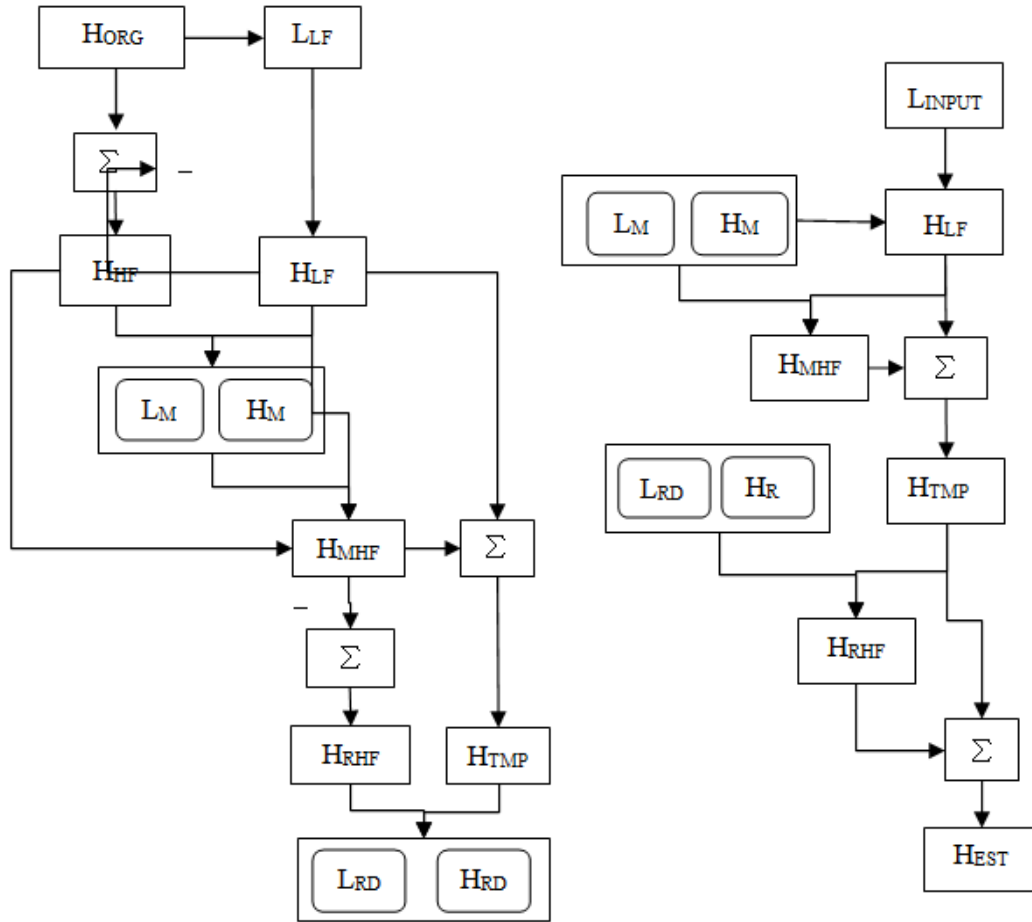
Figure 1. Process of Dictionary Learning and Process of Image Reconstruction

A. Dictionary Training

Firstly, collect a set of training HR images. As illustrated in Figure 1, a HR training image denoted by H_{ORG}, and a corresponding low-resolution image denoted by L_{LF} is constructed using blurred and down-sampled operator. The HR low-frequency image denoted by H_{LF} is generated by applying bicubic interpolation on L_{LF}, as the low-resolution image of the first dictionary training, which is of the same size as H_{org}. Then , by subtracting H_{LF} from H_{org}, HR high-frequency image H_{HF} is constructed as the high-resolution image of the first dictionary training. With the extraction of features from local patches of H_{LF} and H_{HF}, training data $TD = \{p_h^k, p_l^k\}_k$ is constructed. p_h^k is the set of patches extracted from the high-resolution image directly, and p_l^k is the patches built by extracting patches from filtered images obtained by filtering image H_{LF} with certain high-pass filters as described in equation (1). Now, the dimension of low super-resolution image features vectors is 4 times larger than high-resolution image features vectors, so in order to reduce the complexity, reducing the dimensions by Principal Component Analysis (PCA) algorithm.

$$f_1 = [0,0,1,0,0,-1], \quad f_2 = f_1^T, \quad f_3 = [1,0,0,-2,0,0,1], \quad f_4 = f_3^T \tag{1}$$

The main dictionary will be built by training the high-resolution images and

low-resolution images, which is consist of two sub-dictionaries: low-frequency main dictionary L_{MD} and high-frequency main dictionary H_{MD}. Firstly, the K-SVD dictionary training [8] is applied to the set of patches $\{p_l^k\}_k$, generating L_{MD}.

$$L_{MD}, \{q^k\} = \underset{L_{MD}, \{q^k\}}{\operatorname{argmin}} \sum_k \left\| p_l^k - L_{MD} \cdot q^k \right\|_2^2 \qquad s.t. \left\| q^k \right\|_0 \le L \qquad \forall k \qquad (2)$$

Where $\{q^k\}_k$ are sparse representation vectors, and $\|\bullet\|_0$ is L_0 norm counting the nonzero entries of a vector. Based on the assumption that the patch p_h^k can be recovered by approximation as $p_h^k \approx H_{MD} \cdot q^k$, H_{MD} can be defined by minimizing by following mean approximation error.

$$H_{MD} = \underset{H_{MD}}{\operatorname{argmin}} \sum_k \left\| p_h^k - H_{MD} \cdot q^k \right\|_2^2 = \underset{H_{MD}}{\operatorname{argmin}} \sum_k \left\| P_h - H_{MD} \cdot Q \right\|_2^2 \qquad (3)$$

Where the matrix P_h and Q is the set of $\{p_h^k\}_k$ and $\{q^k\}_k$ as its columns, respectively. The solution of the problem is given by the following Pseudo-Inverse expression (given that Q has full row rank):

$$H_{MD} = P_h Q^+ = P_h Q^T (Q Q^T)^{-1} \qquad (4)$$

Finally, the residual dictionary will be trained similar to the training process of main dictionary. The main high-frequency image denoted by H_{MHF} is produced by the image reconstruction of HR low-frequency image H_{LF} with the main dictionary. Then get the input image of residual dictionary learning with the summing of H_{MHF} and H_{LF} denoted by H_{TMP}. As the H_{TMP} which is the low-resolution image of the second dictionary learning stage contains more details than H_{LF} which is the input image of the first dictionary learning stage, and the H_{RHF} is the high-resolution image of the second dictionary learning stage which is deserved by subtracting H_{MHF} from H_{HF}, thus the precise of dictionary learning can be improved.

B. Image Reconstruction

To begin with, a test low-resolution image is given by the same blur and down-sampled operator to original high-resolution image as in dictionary learning phase. Then it is interpolated to the size of original high-resolution image denoted by H_{LF}. The image H_{LF} is filtered by the filters as shown in equation (1) that were used for feature extraction in the training, and patches are extracted from these images which forms the set $\{p_l^k\}_k$.

Combine the high-resolution low-frequency image H_{LF} with the main dictionary MD, the first estimated high-resolution image denoted by H_{MHF} is reconstructed with the following problem which is regularized using the L_1-norm of the coefficients[7].

1) The image H_{LF} is filtered by the filters as shown in equation (1) that were used for feature extraction in the training, and patches are extracted from these images which forms the set $\{p_l^k\}_k$

2) combined patches $\{p_l^k\}_k$ with main dictionary, using OMP algorithm to find the sparse representation vectors $\{q^k\}_k$.

3) the representation vectors $\{q^k\}_k$ are multiplied by H_{MD}, the corresponding high-resolution patches $\{\hat{p}_h^k\}_k \approx \{H_{MD} \cdot q^k\}_k$, are obtained.

Define the operator R_k, which extracts a patch of size n×n from the high-resolution image in location k. then the main high-frequency image denoted by H_{MHF} can calculated by

$$H_{MHF} = \underset{H_{MD}}{\arg\min} \sum_k \left\| R_k H_{MHF} - \hat{p}_h^k \right\|_2^2 \tag{5}$$

This problem has a closed-form Least-Squares solution, given by

$$H_{MHF} = [\sum_k R_k^T R_k]^{-1} \sum_k R_k^T \hat{p}_h^k \tag{6}$$

Then take the H_{MHF} as input and reconstructed the final high-resolution image with the same method.

2.2. Collaborative Representation

As shown in above method, usually regularized the reconstruction problem using the L_1-norm of the coefficients, which is computationally demanding. We can reformulate the problem as a least squares regression regularized by the L2-norm of the coefficients. Thus, we use Collaborative representation and have a closed-form solution. The problem becomes

$$\min_\alpha \left\{ \left\| H_{LF} - L_{MD}\alpha \right\|_2^2 + \lambda \left\| \alpha \right\|_2 \right\} \tag{7}$$

where L_{MD} is the low-resolution dictionary and the parameter λ allows us to alleviate the ill-posed problems and stabilized the solution, which is the coefficients vector. The algebraic solution is given by

$$\alpha = (L_{MD}{}^T L_{MD} + \lambda I)^{-1} L_{MD}{}^T H_{LF} \tag{8}$$

Then the high-resolution patches can be computed using the same coefficients on the high resolution dictionary H_{MD}

$$x = H_{MD}\alpha \tag{9}$$

where x is the HR output patch and H_{MD} the HR dictionary corresponding to L_{MD}. From the equation (8) and equation (9), we obtain

$$x = H_{MD}(L_{MD}{}^T L_{MD} + \lambda I)^{-1} L_{MD}{}^T H_{LF} \tag{10}$$

$$P_G = H_{MD}(L_{MD}{}^T L_{MD} + \lambda I)^{-1} L_{MD}{}^T \tag{11}$$

where P_G is a projection matrix can be computed offline. This means that we only need to multiply the pre-computed projection matrix P_G with the LR input feature

vector, L_{MD}, to calculate the HR output patches X. This can be reduced the computation and complexity greatly during the SR reconstruction algorithm.

At last, the final estimated high-resolution image is constructed with residual dictionary by the same method to the reconstruction with main dictionary. And the input image is the reconstruction image in the first reconstruction phase. We named this algorithm TDL-CR (Two-Layer Dictionary Learning method with Collaborative Representation, TDL-CR).

2.3. Image Reconstruction Based on Svd

Owing to the spatial redundancy, many images do not have full rank. Singular Value Decomposition(SVD) as the important method of matrix decomposition, to compressed large data using the properties of rank deficient of images. For an image x, can be expressed as a matrix, it can be defined as the SVD problem.

$$X = US_X V'$$ (12)

where both U and V are orthonormal matrices and S_X is a diagonal matrix. Equation (12) can also be transformed to a problem of minimize the rank of an reconstruction image.

$$\mathrm{minimize:} \ rank(X) \quad s.t. \|y - Fx\|_2 < \varepsilon$$ (13)

where F is the blur and subsampled operator. Equation (13) can be reformulated to equation (14).

$$\min \|X\|_* \ s.t. \|y - Fx\|_2 \le \varepsilon$$ (14)

where $\|X\|_*$ is nuclear norm of X, defined as sum of singular values. Usually, l_p is a better approximation of the matrix rank than nuclear norm l_1, so we use l_p to replace l_1.

$$\min \|X\|_p, 0 < p < 1 \quad s.t. \|y - Fx\|_2 \le \varepsilon$$ (15)

where $\|X\|_{p^*} = \sum \sigma_i^p$.

The problem is to solve (14). However, it is difficult to solve the constrained problem directly. Therefore we propose to solve the following unconstrained Lagrange version of (14) which is easier to solve

$$J(x) = \min \|y - Fx\|_2 + \lambda \|X\|_p$$ (16)

3. Proposed Methods

Compared with the two-layer dictionary learning method with Collaborative Representation algorithm, the iterative algorithm based on SVD provides more image details with improved PSNR, but the iterative algorithm demand much computation and complexity, thus need much more time to reconstruct image. So we propose an adaptive super-resolution image reconstruction algorithm based on image blocks.

First, divide the image and compute every image block's variance, then classification according to the variance for image blocks of an image, if the variance of image block less than the set threshold, then reconstruct the image area of smoothing and gray changes
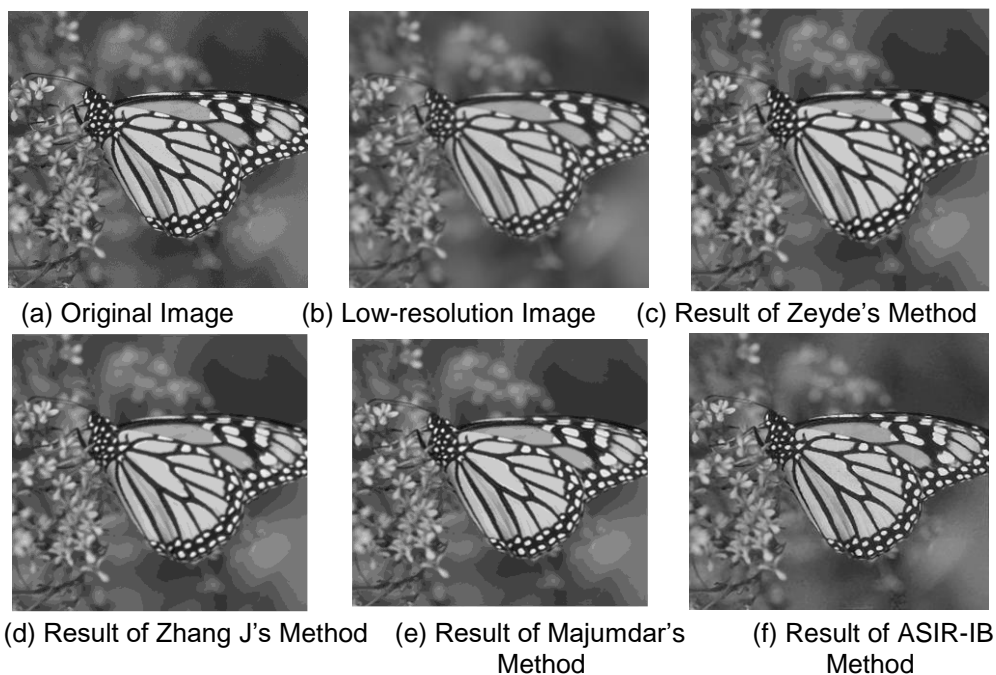
slowly with TDL-CR algorithm, otherwise judgment the image block be the relatively complex block structure, has rich details and reconstruct it with iterative algorithm based on SVD. And then combine all the reconstructed high-resolution image blocks. The global reconstruction constraint enforced by iterative back-projection method to optimization the reconstructed high-resolution image. As we take advantage of TDL-CR method which have the high arithmetic speed, and image reconstruction method of singular value decomposition which have the high precision of image reconstruction, the proposed method can reduced the reconstruction time greatly and barely reduced reconstruction precision.
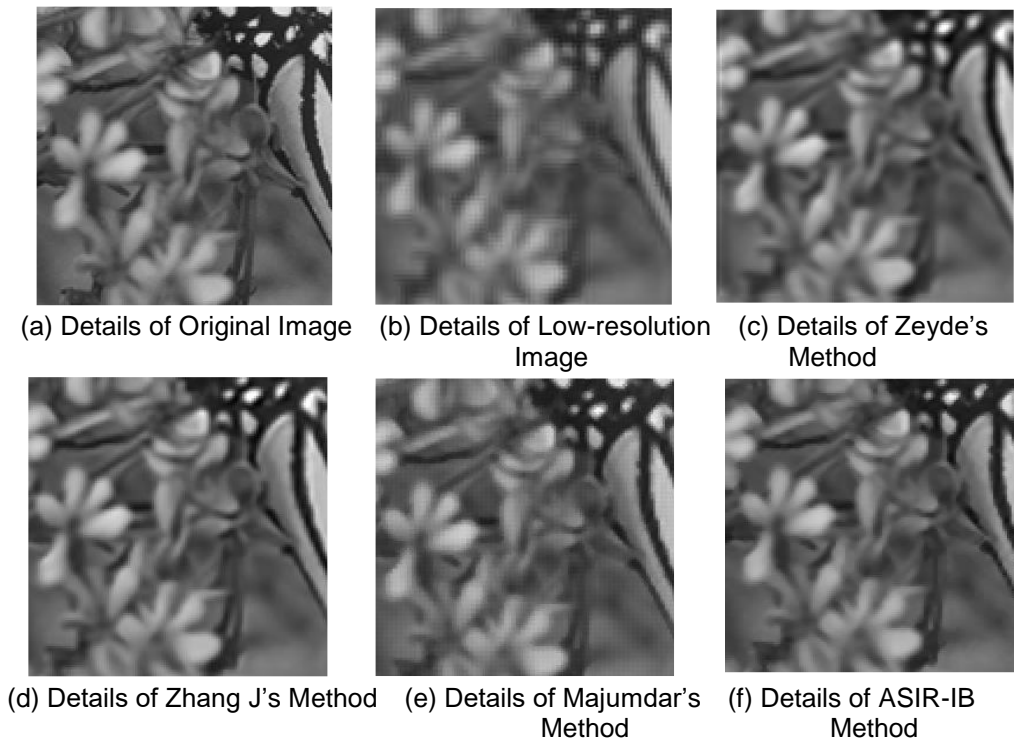
4. Experimental Results

In this section, we first demonstrate the SR results obtained by applying different methods on generic images to identify the effects of our algorithm. We then moved on to discuss various influential factors for the proposed algorithm including the set variance threshold, the size of image block size, size of image patch and different magnify factors.

In our experiments, we magnify the input low-resolution image by a factor of 2 for generic images. We implement our method on some test images to compare with other interpolation methods, the training parameters set as the following.

The blurring operator is 7×7 Gaussian filter with standard deviation of 0.9, the size of MD and the size of RD are both set to 500. The size of image patch is 9×9 with overlap of 3 pixel. The PSNR and time consuming comparing with different methods are listed in Table 1. It can be seen from Table 1 that the proposed method achieves faster time than reference [11] and have the PSNR slightly less than [11]. For up-scaling factor 2, our proposed algorithm is 3 times faster than reference [11]. Due to the image information statics, most of image information are smooth and low-frequency information, high-frequency information and details are only accounts for a small part of image, as the image be partitioned, most smooth area using the two-layer dictionary method to reconstructed high-resolution image block, which greatly reducing image reconstruction time. The visual reconstruction results with different method are presented in Figure 2 and the details of test images are shown as Figure 3.

(a) Original Image   (b) Low-resolution Image   (c) Result of Zeyde's Method

(d) Result of Zhang J's Method   (e) Result of Majumdar's Method   (f) Result of ASIR-IB Method

Figure 2. Reconstructed Images of Monarch

(a) Details of Original Image   (b) Details of Low-resolution Image   (c) Details of Zeyde's Method

(d) Details of Zhang J's Method   (e) Details of Majumdar's Method   (f) Details of ASIR-IB Method

Figure 3. Details of Reconstructed Images of Monarch

Table 1. Comparison of Time Consuming and PSNR for Four Methods

| Method	Zeyde		Zhang J		Majumdar		ASIR-IB	
	PSNR (dB)	Time (s)	PSNR (dB)	Time (s)	PSNR (dB)	Time (s)	PSNR (dB)	Time (s)
Monarch	26.38	3.90	26.73	12.64	28.14	41.82	27.90	14.91
Peppers	30.60	4.13	31.53	12.44	32.51	45.89	31.50	17.00
Lena	29.88	4.20	29.97	12.76	32.23	47.50	31.77	13.44
Cameraman	23.58	1.01	23.78	3.14	24.94	9.01	24.27	5.33
Plane	29.02	4.34	29.23	13.32	31.21	48.72	30.70	14.94
House	29.41	1.06	29.46	3.35	31.48	9.21	30.23	3.17
Average	28.15	3.11	28.45	9.61	30.09	33.69	29.40	11.47

4.1. Effect of the Set Variance Threshold

The previously mentioned experimental results show that the effective of our method to speed up reconstruction process. In those results, we fix the threshold σ to be 35. Intuitively, if the threshold σ to be 0, that means the high-resolution images are obtained by reconstruction algorithm of based on SVD. With the increase of threshold σ, more low-resolution image blocks reconstruct corresponding high-resolution image blocks with two-layer dictionary learning with TDL-CR algorithm, that will lead to the PSNR of the reconstruction image will be decrease, and the speed of reconstruction will be accelerated. Figure 5 shows the PSNR of reconstructed results and the computation time in seconds for test images. We can see that when the threshold σ is 30, the PSNR of reconstructed results will decrease slowly, while the computation time of reconstructed

will decrease fast at 30 and slowly steady at 50, so we set the threshold σ is 30–50.
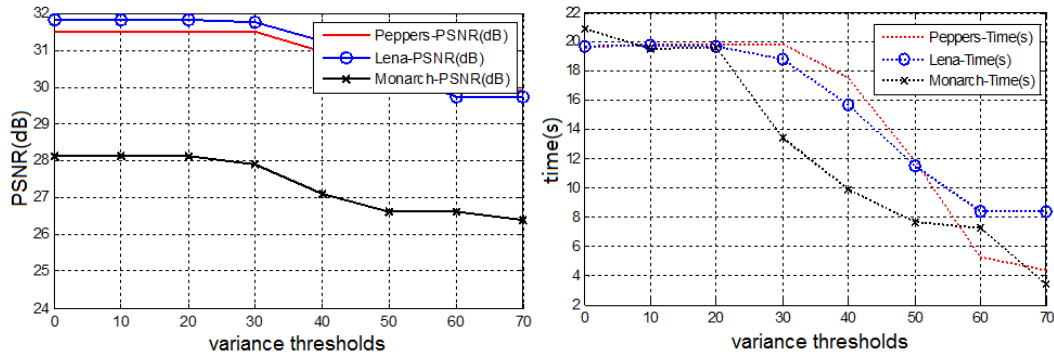
Figure 4. PSNR and Time-Consuming of Reconstructed Results for Different Variance Thresholds

4.2. Effect of Image Block Size

The algorithm need to divided into some blocks thus can adaptive choose different algorithm to reconstructed corresponding high-resolution blocks. While the block size is too small, much blocks need to be reconstructed, thus will lead to too much consuming time to obtain the high-resolution image. In order to analysis the influence of image block size on reconstruction results, we have designed experiments to test the performance of image reconstruction algorithm on different image block size. Figure 6 shows the PSNR of reconstructed results and the computation time in seconds with test images. It can be seen that the PSNR of reconstructed image improved with the increase of block size and steady at block size is 64×64, while less time of image reconstruction needed when block size is 16×16、32×32 and 64×64. We set the block size 64×64 or 32×32 to get the better performance of our algorithm.
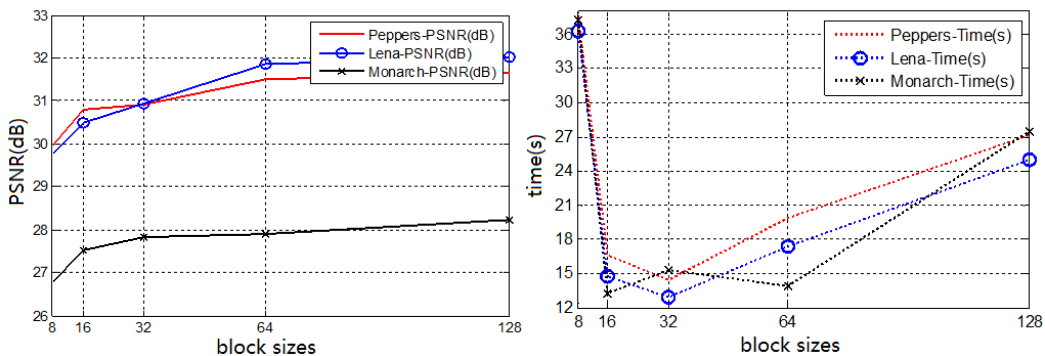

Figure 5. PSNR and Time-Consuming of Constructed Results for Different Block Sizes

4.3. Effect of Image Patch Size in Dictionary Learning

Another important issue of the proposed method is the size of image patch in process of TDL-CR algorithm. Obviously, the patch size cannot be big, otherwise, they will not be micro-structures and hence cannot be represented by a small number of atoms. To evaluate the effects of the patch size on image reconstructed results, we trained the two couple dictionaries and reconstructed high-resolution images with different patch size, 5 $\times 5, 7 \times 7$ and 9×9. Then we applied the two couple dictionaries to the test images. The experimental results of super-resolution are presented in Figure 7, from which we can see that the smaller patch sizes tend to generate some artifacts in smooth regions. Therefore,

we adopt 9×9 as the image patch size in the dictionary learning and reconstruction of our algorithm.

| (5×5) | (7×7) | (9×9) |

| (5×5) | (7×7) | (9×9) |
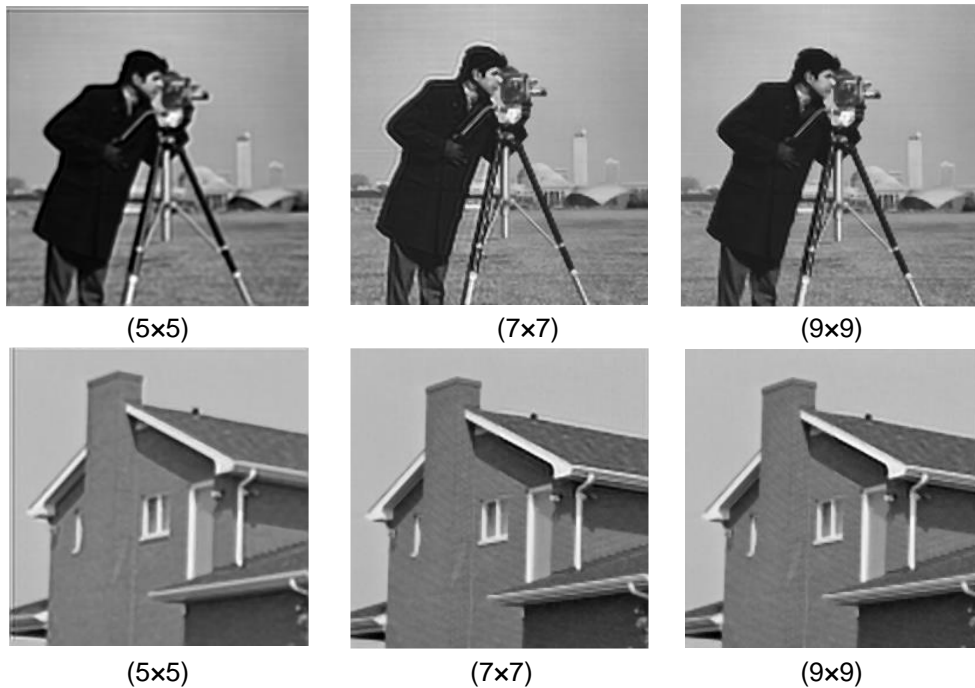
Figure 6. Reconstructed Results of Different Patch Sizes in Dictionary Training Method

4.4. Different Magnifications

To more comprehensively test the adaptability to high-resolution image reconstruction of the proposed image reconstruction method, we performed different up-scaling factors on test images, as shown in Figure 8. The performance in terms of PSNR and running time are shown at Table 2. The larger up-scaling factor tend to lower PSNR and running time, but the proposed method still have better visual reconstructed results, as shown in Figure 9.

(a) Down-sampling of 2   (b) Down-sampling of 3   (c) Down-sampling of 4

Figure 7. Low-Resolution Image at Different Magnifications

| (a) Magnification of 2 | (b) Magnification of 3 | (c) Magnification of 4 |

Figure 8. Results of Image Super-resolution Reconstruction at different Magnifications

Table 2. PSNR and Time Consuming for Reconstruction at Different Magnifications

Magnification	2		3		4	
Test image	PSNR(dB)	Time(s)	PSNR(dB)	Time(s)	PSNR(dB)	Time(s)
Boats	30.16	6.01	24.75	5.31	23.20	4.65
Barbara	27.41	6.32	23.08	5.43	21.64	4.47

5. Conclusions

We proposed an adaptive of image reconstruction method based on image blocks. We also introduced collaborative representation to replace the sparse representation in reconstruction phase of two-layer dictionary learning method, which focuses purely on high arithmetic speed in exchange for some visual quality loss. We combined the advantage of TDL-CR method which have the high arithmetic speed, and image reconstruction method of singular value decomposition which have the high precision of image reconstruction, the proposed adaptive of image reconstruction method can greatly decrease the image reconstruction time and barely reduce the reconstruction precision.

References

[1] W. T. Freeman, E. C. Pasztor, O. T. Carmichael, "Learning low-level vision," International Journal of Computer Vision, vol. 40, no. 1, **(2000)**, pp. 25–47,

[2] J. Sun, N. N. Zheng, H. Tao, and H. Shum, "Image hallucination with primal sketch priors," IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, **(2003)**, pp. 729–736.

[3] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," IEEE Conference on Computer Vision and Pattern Classification, vol. 1, **(2004)**, pp. 275–282.

[4] Z. Xiong, X. Sun, and F. Wu, "Image hallucination with feature enhancement," IEEE Conference on Computer Vision and Pattern Classification, vol. 1, **(2009)**, pp. 2074–2081.

[5] D. L. Donoho, "Compressed sensing," IEEE Transactions on Information Theory, vol. 52, no. 4, **(2006)**, pp. 1289–1306.

[6] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE Trans. on Image Processing, vol. 19, no. 11, **(2010)**, pp. 2861–2873.

[7] R. Zeyde, M. Elad, and M. Protter. "On single image scale-up using sparse-representations," InCurves and Surfaces, **(2010)**, pp. 711–730.

[8] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Trans. on Signal Processing, vol. 54, no. 11, **(2006)**, pp. 4311–4322.

[9] M. Elad, M. Aharon, "Image denoising via learned dictionaries and sparse representation," International Conference on Computer Vision and Pattern Recognition, New York, **(2006)**, pp. 17-22.

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

[11] A. Majumdar, R K. Ward "An algorithm for sparse MRI reconstruction by schatten p-norm minimization," Magnetic Resonance Imaging, vol. 29, no. 3, **(2011)**, pp. 408-417.

Authors

Yong Yin, he received the Ph.D degree in control engineering from Chongqing University, China, in 1999. In 1984, he joined College of Communication Engineering, Chongqing University, China, where he was engaged in research on digital image and video processing. He is a professor in Chongqing University. His current research interests are centered around image processing and recognition and image and video retrieval.

Qianqian Ruan, she received her B. Sc. Degree in Communication from Luoyang Normal University, China in 2013, and received the B.S. degree in Communication from Choingqing University, China in 2016. where herresearch interests include pattern recognition and computer vision.

Shitong Liu, he received his B. Sc. Degree in Communication from Chongqing University of Posts and Telecommunications, China in 2015, He is currently a master candidate of Chongqing University, China, and his research interests include machine learning and image processing.

Xiaoli Chen, she received her B. Sc. Degree from Northwestern Polytechnical University, China in 2014, She is currently a master candidate of Chongqing University, China, where her research interests include pattern recognition and computer vision.