Image Super-Resolution Reconstruction Based on Two-Stage Dictionary Learning

Li Shang[1] and Zhan-li Sun[2]

[1] Department of Communication Technology,
College of Electronic Information Engineering,
Suzhou Vocational University, Suzhou 215104, Jiangsu, China
[2] Department of Automation, College of Electrical Engineering and Automation,
Anhui University, Hefei 230601, Anhui, China
sl0930@jssvc.edu.cn, zhlsun2006@yahoo.com.cn

Abstract. The general image super-resolution reconstruction (SRR) methods based on sparse representation utilizes the one-stage high and low resolution dictionary pairs to reconstruct a high resolution image, and this method can not restore much image detail information. To solve this detect, two-stage high and low resolution dictionaries are explored here. The goal of exploiting the two-stage dictionaries is to reconstruct the difference image between the original high resolution image and the reconstructed image obtained by using the one-stage dictionaries. In learning two-stage dictionaries, the difference image is used as the high resolution (HR) image, and the first-order and second-order gradient feature images of the one-stage reconstructed images are used as the low resolution (LR) images. Then, the two-stage dictionaries are learned by K-singular value decomposition (K-SVD) method. In test, an artificial and a real LR image are used, and simulation results show that, compared with other learning-based methods, our method proposed has remarkable improvement in PSNR and visual effect.

Keywords: Super-resolution reconstruction (SRR), Sparse representation, K- singular value decomposition (K-SVD), One-stage dictionaries, Two-stage dictionaries, Low resolution, High resolution.

1 Introduction

At present, image super-resolution reconstruction (SRR)methods based on learning have been the research hot [1-2]. The latest learning based method is that proposed by Yang et al [3-4] denoted SC-SRR method, here SC is the abbreviation of sparse coding. The theory of SC-SRR is that the coefficients of high resolution (HR) images can be represented by using those of low resolution (LR) images [5-7]. This method can restore certain high frequency details of images, but the reconstructed results have certain block effect to some extent. To restore much image details and improve the quality of reconstructed images, on the basis of SC-SRR, the two-stage HR and LR dictionary learning is proposed here. The one-stage HR and LR dictionaries are

D.-S. Huang et al. (Eds.): ICIC 2014, LNAI 8589, pp. 277–284, 2014.

learned by SC-SRR method. Utilizing the one-stage reconstructed image as the HR features, as well as the first order and the second order gradient features of the one-stage reconstructed image as LR features, the two-stage HR and LR dictionaries can be learned by the K- singular value decomposition (K-SVD) method [8-10]. In test, an artificial and a real millimeter wave (MMW) image are utilized. Compared with other learning-based methods, simulation results show that our method has great improvement in image structural similarity and visual effect.

2 Sparse Representation Based SRR Method

Let the matrix $D \in \Re^{N\times K}$ （$n << K$） be an over-complete dictionary of K prototype atoms, and supposed a vector $x \in \Re^{N}$ can be represented as a sparse linear combination of these atoms. Thus, this signal x can be approximately written as

$$\min_{s} \|s\|_{p}, \quad s.t. \quad \|x-Ds\|_2^2 \le \varepsilon \quad . \tag{1}$$

where $s \in \Re^{K}$ represents the sparse coefficient vector [5], $\|\cdot\|_p$ and $\|\cdot\|_2$ denote respectively the l_p and l_2 norm. Assumed that x_h and x_l denote respectively a HR and LR image patch vector, according to Eqn. (1), x_h can be approximated by the equation of $x_h \approx D_h s$, and x_l can be approximated by $x_l = T x_h \approx T D_h s$, here T is the mapping matrix. Thus, the LR dictionary D_l can be calculated by $D_l = T D_h$. Namely, the HR image patch x_h can be restored by using the equation of $x_h = D_h s$. Commonly, for an image I , to improve the quality of reconstructed images, it is randomly sampled L times with $p \times p$ patch, denoted by matrix $X = \{x_1, x_2, \cdots, x_L\} \in \Re^{N\times L}$, thus, the dictionary D is learned from X , and the optimized problem is described as

$$\{D,S\} = \arg\min_{D,S} \|X - DS\|_F^2 + \lambda \|S\|_1 \quad . \tag{2}$$

subject to $\|d_k\|_2^2 \le 1$ ($k = 1,2,3,\cdots,K$), d_k is the kth atom of the dictionary D , and S denotes the sparse coefficient matrix.

3 Training Two-Stage Dictionaries

3.1 SC- SRR Based One-Stage Dictionary Pairs Learning

The one-stage HR and LR dictionary pairs are trained here by Yang's SC-SRR method [4-5]. This problem by generalizing the basic SC scheme as follows [4-5]

$$\min_{D_h, D_l, \left\{s_i^{x_h/x_l}\right\}_{i=1}^{L}} \sum_{i=1}^{L} \left\{ \|x_{hi} - D_h s_i\|_2^2 + \|x_{li} - D_l s_i\|_2^2 \right\} + \lambda \|s_i\|_1 \quad . \tag{3}$$

Further, to guarantee the compatibility between adjacent patches, Yang et al modified the optimization problem of Equation (3), which is written as follows [6-7]

$$\min_{S}\left\{\frac{1}{2}\left\|\begin{bmatrix}Fx_{li}\\ \lambda x_{hi}\end{bmatrix}-\begin{bmatrix}FD_l\\ \lambda PD_h\end{bmatrix}\cdot s_i\right\|_2^2+\lambda\|s_i\|_1\right\}=\left\{\frac{1}{2}\left\|\tilde{y}-\tilde{D}s_i\right\|_2^2+\lambda\|s_i\|_1\right\}. \tag{4}$$

Where Eqns. (3) and (4) are subjected to $\|D_h(:,k)\|_2\le 1$ and $\|D_l(:,k)\|_2\le 1$ ($k=1,2,3,\cdots,K$). Parameter P extracts the region of overlap between current target patch and previously reconstructed high-resolution image, and F is a feature extraction operator. Given the optimal solution s^*, the HR patch can be reconstructed as $\hat{x}_h=D_h s^*$, and the patch $\hat{x}_h$ is put into the HR image $\hat{X}_{h1}$ set.

3.2 Two-Stage Dictionary Pairs Learning

Assumed that the original HR image is denoted by I_h, and it can be represented as $I_h=\hat{I}_1+I_e$, where I_e is the difference image between the HR image I_h and the one-stage reconstructed image $\hat{I}_1$. The goal of training two-stage high and LR dictionaries is to reconstruct the difference image and make it appreciate I_e. Here, the training of two-stage dictionaries are mainly generalized as follows:

Step 1. Ascertaining the HR image patch set X^{h2}. The difference image I_e is used as the HR image. Randomly sampled I_e image L times with $d\times d$ pixels, the HR image patch set $X^{h2}=\{x_1^{h2},x_2^{h2},\cdots,x_L^{h2}\}$ with the size of $d^2\times L$ can be obtained.

Step 2. Ascertaining the LR image patch set X^{l2}. Because the high frequency information of LR images has important effect on predicting detail information of HR images, for the one-stage reconstructed image $\hat{I}_1$, four one dimension gradient filters [19] are used to improve the prediction precision of HR images' details. Utilizing four gradient filters, the first order and second order gradient information of $\hat{I}_1$, distributing in vertical and horizontal direction, can be obtained. Therefore, the four gradient feature images obtained are used as the LR features of the two-stage LR dictionary. For each gradient feature image, the image patch random sampling method is the same as that utilized in Step 1 so as to obtain a LR image patch set with $d^2\times L$ size, denoted by $X_{rg}^{l2}=\{x_{rg1}^{l2},x_{rg2}^{l2},\cdots,x_{rgL}^{l2}\}$ $(r=1,2,3,4)$, which represents respectively the first order horizontal and vertical gradient feature image patch set (i.e., X_{1g}^{l2} and X_{2g}^{l2} set), and the second order horizontal and vertical gradient feature image patch set (i.e., X_{3g}^{l2} and X_{4g}^{l2} set), . And then, connected each gradient feature image patch set in order, the total LR image patch training set $X^{l2}=\{X_{1g}^{l2};X_{2g}^{l2};X_{3g}^{l2};X_{4g}^{l2}\}$ can be obtained.

Step 3. Selecting the high frequency features. Because the human visual system (HVS) is sensitive to image edge features, the high frequency features are selected as cluster features. For an original HR image I_h, its high frequency features are obtained by using contourlet transform method [20-21]. Here, the number of contourlet decomposition layer is set as 2, thus each layer has four high pass sub-band images. For each layer, the fusion image of four high pass sub-band images is first considered, and each layer's fusion result is denoted by I_h^{f1} and I_h^{f2}. And then, the fusion between I_h^{f1} and I_h^{f2} are considered again, and this high pass sub-band images' fusion result between each layer is denoted by I_h^f. Further, I_h^f image is segmented into M image patches by using the fixed $d \times d$ sub-window, and the cluster sample set of high frequency features $X_{hf} = \{x_{hf1}, x_{hf2}, \cdots, x_{hfM}\}$ with $d^2 \times M$ size are obtained.

Step 4. Defining classified marking vector and cluster center of high frequency fusion feature set X_{hf}. Set the threshold Δ, for $\forall x_{hfi} \in X_{hf} = \{x_{hf1}, x_{hf2}, \cdots, x_{hfM}\}$, if the variance of x_{hfi}, denoted by $\mathrm{var}(x_{hfi})$, is smaller than Δ, then the set satisfying the condition of $\mathrm{var}(x_{hfi}) < \Delta$ is defined as the smooth set X_{hfs}, conversely the remaining x_{hfi} elements comprise the feature set X_{hff}. Utilized K-means method, the sample set of K classes and the K cluster centers can be obtained.

Step 5. Classifying the HR and LR set X'^{h2} and X'^{l2}. The class label of the HR and LR vector x_i^{h2} and x_i^{l2} are determined by ϕ_i. Thus, the $K+1$ class HR and LR sample set can be obtained, denoted respectively by $X'^{h2} = \{X'^{h2}_1, X'^{h2}_2, \cdots, X'^{h2}_K, X'^{h2}_{K+1}\}$ and $X'^{l2} = \{X'^{l2}_1, X'^{l2}_2, \cdots, X'^{l2}_K, X'^{l2}_{K+1}\}$.

Step 6. Training the two-stage HR and LR dictionary pairs. The two-stage LR dictionary D^{l2} corresponding to X'^{l2} is trained by K-SVD algorithm. Utilized the LR coefficient matrix S_K and the HR training set X'^{h2}, the two-stage HR dictionary D_k^{h2} can be learned by $D_k^{h2} = \arg\min \left\| X'^{h2}_k - D_k^{h2} S_k \right\|_F^2$. Combined D_k^{h2} and D_k^{l2}, the two-stage dictionary pairs $\{D_k^{h2}, D_k^{l2}\}$ can be obtained.

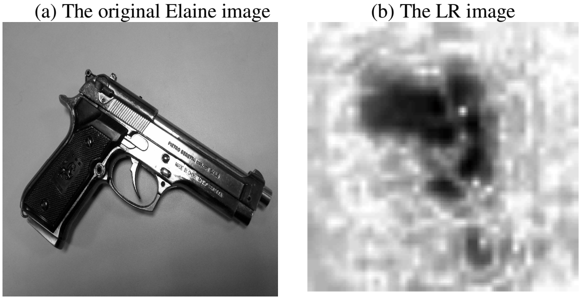
4 Experimental Results and Analysis

4.1 SRR Results Using One-Stage Dictionary Pairs

In test, five natural images with 512×512 size were used. Each original image was sampled randomly with 8×8 image patch 10000 times, thus the HR image patch set denoted by X^{h1} with 64×50000 size was obtained. For each HR image, using a point spread function (PSF) filter and the down-sampling method, the corresponding LR image can be obtained. For example, a test image called Elaine and its LR version was shown in Fig. 1(a) and Fig.1(b). At the same time, the real LR image, namely the

MMW image, was also shown in Fig.1(d). Using the same image patch sampling method, the LR image patch set denoted by X^{l1} with 64×50000 size could be obtained. Let X^{h1} and X^{l1} to be the training samples of SC-SRR, it can be obtained the one-stage HR and LR dictionary pairs, denoted by D^{h1} and D^{l1} with 256 atoms.

(a) The original Elaine image (b) The LR image



(c) The original toy gun image (d) MMW image

Fig. 1. The original images and corresponding low-resolution versions

And for given LR dictionary D^{l1} and the LR image patch feature vector $\{x_i^{l1}\}$, the LR coefficient vector $\{s_i^{l1}\}$ can be learned by the following form

$$s_i^{l1} = \arg\min_s \frac{1}{2} \left\| x_i^{l1} - D^{l1} s_i^{l1} \right\|_2^2 + \gamma \left\| s_i^{l1} \right\|_1 . \tag{5}$$

Utilizing the LR coefficient s_i^{l1} and the HR dictionary D^{h1}, the reconstructed HR image patch $\hat{x}_i^{h1}$ can be solved by $\hat{x}_i^{h1} = D^{h1} s_i^{l1}$, further, replaced $\hat{x}_i^{h1}$ to the corresponding location in $\hat{I}_1$, $\hat{I}_1$ can be reconstructed. Here, considering the paper's length, only the reconstructed examples of Elaine LR version and MMW image were given shown in Fig.2. Clearly, LR images' contour has been reconstructed, however, LR images' many details are not still restored, at the same time, the restored image edge is very blurred from visual appeal. Especially, for the MMW image restored by the one-stage reconstructed process, the background noise is not reduced well.

(a) Elaine image (b) MMW image

Fig. 2. SRR images obtained by using one-stage high and LR dictionary pairs. (a) Reconstructed Elaine image. (b) Reconstructed MMW image.

4.2 SRR Results Using Two-Stage Dictionary Pairs

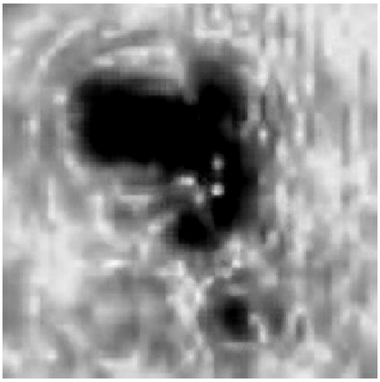
In this stage, for each reconstructed image $\hat{I}_1$, its high frequency feature set $\{x_{hfi}\}$ can be obtained by contourlet transform with 2 layers and 4 directions, then the class of x_{hfi} can be determined by K-means cluster method. And the four order gradient images of $\hat{I}_1$ are used as the LR feature set $\{x_i^{l2}\}$, and its class is defined by that of $\{x_{hfi}\}$. Then, according to the threshold Δ, the high frequency feature set $\{x_{hfi}\}$ is divided into the feature set $\{x_{hffi}\}$ and the smooth set $\{x_{hfsi}\}$. The class of image patches in the set $\{x_{hfsi}\}$ is marked as the $K+1$ class, and the corresponding LR feature image patch can be represented by LR dictionary D_{K+1}^{l2}. For feature set $\{x_{hfi}\}$, the class is defined by the form of $k_i = \arg\min_k \left\| x_{hfi} - c_k \right\|_2$. Thus, the set x_i^{l2} can be represented by using the $k_i - th$ dictionary $D_{k_i}^{l2}$. Further, the LR sparse coefficient vector s_i^{l2} can be obtained by the equation of $s_i^{l2} = \arg\min_s \frac{1}{2}\left\| x_i^{l2} - D_{k_i}^{h2} s_i^{l2} \right\|_2^2 + \gamma \left\| s_i^{l2} \right\|_1$., the two-stage reconstructed image $\hat{I}_2$ can be computed by the following form:

$$\hat{I}_2 = \hat{I}_1 + \left[\sum_i (R_i)^T R_i \right]^{-1} \left[\sum_i (R_i)^T D_{k_i}^{h2} s_i^{l2} \right] . \quad (6)$$

where R_i is the ith image patch of the extracted image, and it is converted a column vector in solving the optimization of Eqn.(6).

And utilized Eqn. (6), the two-stage reconstructed images could be obtained. In the same way, considering the paper's length, only the reconstructed results of Elaine LR image and MMW image were given as shown in Fig.5 by using our method.In the same way, the reconstructed results of Elaine LR image and MMW image were given

(a) Elaine image



(b) MMW image

Fig. 3. SRR images obtained by using the two-stage high and LR dictionary pairs. (a) Reconstructed Elaine image. (b) Reconstructed MMW image.

as shown in Fig.3 by our method. Compared Fig.3 and Fig.2, from the visual effect, the edge of the reconstructed Elaine image shown in Fig.3 (a) is clearer than those shown in Fig.2 (a), and much detail information is restored. And for reconstructed MMW image, much background noise in Fig.3 (b) has been reduced greatly.

In order to further differentiate the performance of different SRR algorithms, for reconstructed natural images, the PSNR measurement criterion is used to estimate the equality of reconstructed results. For the reconstructed MMW image, it is measured by using the relative SNR (RSNR) criterion. The calculated values of PSNR with different algorithms were listed in Table 1. And for the MMW image, the computed RSNR values were shown in Table 2.

Table 1. PSNR values of restored images using different algorithms (LR images were processed by PSF filter with Gaussian variance 2 and downsampling with the extracted scale 3)

Algorithms / Images	Our method	SC-SRR	K-SVD	Noise images
Elaine	20.53	19.26	18.14	4.97
Lena	21.55	19.37	18.21	5.66
Boat	20.78	18.46	17.72	5.34
House	22.62	21.53	19.17	7.83

Table 2. RSNR values of restored MMW images using different algorithms

Algorithms / Images	Our method	SC-SRR	K-SVD	MMW image
Reconstructed MMW image	15.36	14.84	14.63	12.37

5 Conclusions

A novel image SRR method using the two-stage dictionary pairs with class information is discussed in this paper. The one-stage HR and LR dictionary pairs are learned by SC-SRR method. And the two-stage HR and LR dictionaries are trained by K-SVD algorithm. In test, the restoration effect of our method is testified by using five highly degraded natural images and a real MMW image. Experimental results show that, compared with methods of K-SVD and SC-SRR, our method indeed has clear improvement in PSNR for natural images and in RSNR for the MMW image, and has better image structures and visual effect.

Acknowledgement. This work was supported by the grants from National Nature Science Foundation of China (Grant No. 61373098 and 61370109), the Innovative Team Foundation of Suzhou Vocational University (Grant No. 3100125),the grant from Natural Science Foundation of Anhui Province (No. 1308085MF85).

References

1. Elad, M.: Sparse and redundant representation: From theory to applications in signal and image processing, pp. 228–237. Springer, New York (2010)
2. Li, X., Orchard, M.T.: New edge-directed interpolation. IEEE Transactions on Image Processing 10(10), 1521–1527 (2001)
3. Yang, J.C., Wright, J., Ma, Y., Huang, T.: Image super-resolution as sparse representation of raw image patches. In: Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE Press, Anchorage (2008)
4. Yang, J.C., Wright, J., Huang, T., Ma, Y.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing 19(11), 2861–2873 (2010)
5. Zeyde, R., Elad, M., Protter, M.: On single image scale-up using sparse-representations. In: Boissonnat, J.-D., Chenin, P., Cohen, A., Gout, C., Lyche, T., Mazure, M.-L., Schumaker, L. (eds.) Curves and Surfaces 2011. LNCS, vol. 6920, pp. 711–730. Springer, Heidelberg (2012)
6. Freeman, W.T., Pasztor, E.C., Carmichael, O.T.: Example-based super-resolution. IEEE Computer Graphics and Application 22(2), 56–65 (2002)
7. Huang, D.S., Horace, H.S., lp, Law Ken, C.K., Chi, Z.R.: Zeroing polynomials using modified constrained neural network approach. IEEE Trans. on Neural Networks 16(3), 721–732 (2005)
8. Huang, D.S., Horace, H.S., lp, Ken, C.K., Chi, Z.R., Wong, H.S.: A new partitioning neural network model for recursively finding arbitrary roots of higher order arbitrary polynomials. Applied Mathematics and Computation 162(3), 1183–1200 (2005)
9. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. IEEE Transactions on Signal Processing 54(11), 4311–4322 (2006)
10. Huang, D.S., Horace, H.S., lp, et al Zheru, C.: Dilation method for finding close roots of polynomials based on constrained learning neural networks. Physics Letters A 309(5-6), 443–451 (2003)