Journal of Southeast University (English Edition) Vol. 32, No. 2, pp. 177 - 182 June 2016 ISSN 1003—7985

Single frame super-resolution reconstruction based on sparse representation

Xie Chao[1,2] Lu Xiaobo[1,2] Zeng Weili[3]

([1]School of Automation, Southeast University, Nanjing 210096, China)

([2]Key Laboratory of Measurement and Control of Complex Systems of Engineering of Ministry of Education, Southeast University, Nanjing 210096, China)

([3]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: In order to effectively improve the quality of recovered images, a single frame super-resolution reconstruction method based on sparse representation is proposed. The combination method of local orientation estimation-based image patch clustering and principal component analysis is used to obtain a series of geometric dictionaries of different orientations in the dictionary learning process. Subsequently, the dictionary of the nearest orientation is adaptively assigned to each of the input patches that need to be represented in the sparse coding process. Moreover, the consistency of gradients is further incorporated into the basic framework to make more substantial progress in preserving more fine edges and producing sharper results. Two groups of experiments on different types of natural images indicate that the proposed method outperforms some state-of-the-art counterparts in terms of both numerical indicators and visual quality.

Key words: single frame super-resolution reconstruction; sparse representation; local orientation estimation; principal component analysis (PCA); consistency of gradients

doi: 10.3969/j.issn.1003-7985.2016.02.008

Images with high-resolution (HR) are usually desired in most electronic imaging applications such as medical imaging, remote sensing, and video surveillance[1], etc. The most straightforward solutions to an enhancement in spatial resolution are to either reduce the pixel size or increase the image chip size via sensor manufacturing techniques. However, both of them are severely restricted by the physical limitations of the imaging system[2]. Therefore, it is much more promising and economical to achieve the goal of resolution enhancement via image processing techniques.

Single frame super-resolution reconstruction is such a technique, and it will be referred to as "SR" hereafter. SR is cast as the ill-posed inverse problem of recovering an ideal HR image from only one observed low-resolution (LR) image suffering from both blurring and noise. Conventional SR approaches can be mainly divided into two categories: interpolation-based methods[3] and regularization-based methods[4]. Bilinear and bicubic interpolations are the most common interpolation-based methods, which utilize polynomial approximation models to calculate every missing pixel within a local neighborhood. Their advantages are simple to understand and easy to operate, whereas the fatal shortcoming is that they can barely recover the high-frequency components lost or degraded during the high-to-low sampling process. In order to find a kind of better method, researchers turn to regularizing the SR problem via some prior knowledge of natural images, i.e., regularization-based methods. The total variation (TV) model[5-6], using the l_1-norm of the first-order derivative of the HR image as the regularization term, is one of the most commonly used regularization methods and has been widely employed in the fields of image denoising and super-resolution. However, due to the characteristics of TV, it tends to smooth down the delicate details of the output images. Consequently, some improving algorithms based on TV are developed to overcome this weakness[7-8].

In light of the fact that many natural images are actually sparse or compressible when expressed on a proper basis, another category, namely sparse representation-based methods, has already been actively developed and successfully applied in the application of SR[9]. The representative work is proposed by Yang et al.[10-11]. In their papers, a pair consisting of a HR dictionary and its corresponding LR dictionary are learned in advance by extracting raw patches from some training images randomly, and then sparse coding is applied to the overlapping patches sampled in raster-scan order from the input LR image to obtain the sparse coefficients. Finally, the SR output is recovered by averaging the overlapping HR patches produced by the coefficients and HR dictionary. Although

Received 2015-10-18.

Biographies: Xie Chao (1987—), male, graduate; Lu Xiaobo (corresponding author), male, doctor, professor, xblu2013@126.com.

Foundation items: The National Natural Science Foundation of China (No. 61374194, No. 61403081), the National Key Science & Technology Pillar Program of China (No. 2014BAG01B03), the Natural Science Foundation of Jiangsu Province (No. BK20140638), the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Citation: Xie Chao, Lu Xiaobo, Zeng Weili. Single frame super-resolution reconstruction based on sparse representation[J]. Journal of Southeast University (English Edition), 2016, 32(2): 177 - 182. doi: 10.3969/j.issn.1003-7985.2016.02.008.

this scheme is proved to lead to a state-of-the-art result, it is very time-consuming to obtain two large dictionaries generated by random sampling. Hence, in Ref. [12], an improved method is put forward, where a lessened dictionary is utilized to accelerate the dictionary learning process. Clustering techniques are added to the sparse representation-based SR scheme to exploit the priors of the extracted raw patches. Yang et al. [13] employed multiple dictionaries learned from several groups of patches produced by the K-means algorithm, where the superiority to the prototype dictionary is experimentally validated. Peleg et al. [14] also suggested a statistical prediction model which results in better HR coefficients. The HR coefficients are predicted from LR coefficients via a MMSE estimator.

From the analyses above, a conclusion can be drawn that the determination of an efficient dictionary is a critical issue in the sparse representation-based SR scheme. Lu et al. [15] pointed out that the performance of sparse representation can be greatly improved by taking the advantage of the intrinsic geometric structure of training image patches. Hence, it is crucial to explore the potential geometric structure to enhance the existing SR performance. In this paper, we present a geometric structure-based dictionaries learning method, by which the geometric dictionaries are learned. Moreover, the most suitable dictionary is selected from the learned dictionaries for each image patch to be processed. Due to the reason that each part of the image is represented via the well-selected dictionary, a better result can be obtained than merely employing only one universal dictionary. Besides dictionary learning, one additional regularization term can also be introduced into the basic SR scheme to further improve reconstruction performance. The measurement of gradient difference between the reconstructed HR image and the observed LR image is used as the regularization term here, which is significant for preserving the sharpness of the edges in the reconstructed image. After introducing the proposed SR method, an efficient iterative shrinkage algorithm is adopted to mathematically solve it.

1 Proposed Method

Suppose that $\boldsymbol{X}$ is the ideal HR image, while the LR image $\boldsymbol{Y}$ is a blurred and down-sampled version of $\boldsymbol{X}$ in the same scenario. The observation model of the single frame SR can be described as

$$\boldsymbol{Y} = \boldsymbol{QHX} + \boldsymbol{v} \tag{1}$$

where $\boldsymbol{H}$ represents the blurring operator; $\boldsymbol{Q}$ is the down-sampling operator; and $\boldsymbol{v}$ is the error or additive noise. The purpose of SR can be regarded as estimating the ideal $\boldsymbol{X}$ as precisely as possible using only one input $\boldsymbol{Y}$.

Since much information is discarded during the high-to-low image acquisition process, the linear equation (1) is severely underdetermined; i. e., infinite many solutions may be suitable for (1). In order to obtain a proper solution, an additional regularization term is usually introduced into the basic observation model for the reason that this term can bring some priors of natural images. Consequently, SR can be transformed into a more generalized mathematical problem, which minimizes the cost function as

$$\min J(\boldsymbol{X}),\ J(\boldsymbol{X}) = \Psi(\boldsymbol{Y}, \boldsymbol{QHX}) + \upsilon(\boldsymbol{X}) \tag{2}$$

The first term in the right-hand side of (2) is called the data error term, and Ψ is the function to measure the distance between the estimated and the observed images. The second term is called the regularization term (or penalty term).

1.1 Data error term

In most cases, the data error term in (2) is defined by the l_p-norm of the residual, which can be written as

$$\Psi(\boldsymbol{Y}, \boldsymbol{QHX}) = \left\| \boldsymbol{Y} - \boldsymbol{QHX} \right\|_p^p \tag{3}$$

According to Refs. [11 −16], the l_2-norm is more suitable for the framework of sparse representation that is used in this paper. As a result, the l_2-norm is adopted in this paper.

1.2 Sparse representation term

The regularization term in (2) is of great significance, since it makes the underdetermined SR problem solvable and ensures a proper and stable solution. Under the framework of sparse representation, the SR problem is regularized by the sparse prior that every local patch $\boldsymbol{x}_i$ extracted from HR image $\boldsymbol{X}$ can be represented as a linear combination of few atoms (i. e., columns) from a proper dictionary learned from a set of training images. Suppose that $\boldsymbol{X}$ is of size $\sqrt{N} \times \sqrt{N}$, then rearranged lexicographically to obtain a column vector $\boldsymbol{X} \in \mathbf{R}^N$. The $n \times N$ matrices $\boldsymbol{R}_i$ is employed to extract the i-th local patch from $\boldsymbol{X}$, i. e., $\boldsymbol{x}_i = \boldsymbol{R}_i \boldsymbol{X}$. Considering the situation of maximal overlaps, a total number of $M = (\sqrt{N} - \sqrt{n} + 1)^2$ patches can be extracted from $\boldsymbol{X}$. For each patch $\boldsymbol{x}_i$, assume that a known and proper dictionary $\boldsymbol{D}_{t_i}$ is selected. Then the estimate of $\boldsymbol{x}_i$ and its sparse coefficient $\boldsymbol{\alpha}_i$ can be computed by a sparse coding operation. Subsequently, the optimal estimate of the HR image is computed by averaging all the estimated local patches according to Ref. [17].

$$\hat{\boldsymbol{X}} = \left(\sum_{i=1}^{M} \boldsymbol{R}_i^{\mathrm{T}} \boldsymbol{R}_i \right)^{-1} \left(\sum_{i=1}^{M} \boldsymbol{R}_i^{\mathrm{T}} \boldsymbol{D}_{t_i} \boldsymbol{\alpha}_i \right) \tag{4}$$

The matrix to be inverted in Eq. (4) is a diagonal one, and thus the computation can be done on a pixel-by-pixel basis to reduce the burden of computation. Additionally, working on the overlapping patches and then averaging the results can efficiently suppress noise and prevent block

artifacts. Substituting Eqs. (3) and (4) into Eq. (2), the basic objective function is formulated as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}}\{\left\|\boldsymbol{Y}-\boldsymbol{QH}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_i\right)^{-1}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_{t_i}\boldsymbol{\alpha}_i\right)\right\|_2^2+\sum_{i=1}^{M}\|\boldsymbol{W}_i\boldsymbol{\alpha}_i\|_1\} \tag{5}$$

where $\boldsymbol{\alpha}$ is the concatenation of all sparse coefficients $\boldsymbol{\alpha}_i$, and $\boldsymbol{W}_i$ is a diagonal matrix whose diagonal entries are inversely proportional to the corresponding magnitude of sparse coefficients, i. e., $w_{i,j}=1/(\alpha_{i,j}+\varepsilon)$[18]. The sparsity regularization term has already been changed from l_0-norm to l_1-norm as long as the coefficients are sufficiently sparse due to the non-convexity of l_0-norm[19-20].

1.3 Geometric dictionaries learning and selection

Clearly, two key procedures left incomplete in the previous subsection are dictionary learning and selecting one dictionary for each local patch adaptively. Considering that the potential geometric structure is of great significance in enhancing the SR performance, a local orientation estimation-based clustering method[16,21] is used to pre-classify the raw training patches into K clusters. Since the patches belonging to the same cluster have similar orientations, we decide to learn a PCA-based compact dictionary from them instead of an over-complete one, and the final geometric dictionaries are the concatenation of all K compact dictionaries. The specific process of this dictionary learning method is summarized as follows:

Step 1 Extract raw patches from training images randomly, and estimate the local orientations of each patch based on the calculation procedures described in Ref. [21].

Step 2 Classify these patches into K directional clusters according to estimated local orientations.

Step 3 For each cluster, apply PCA to it to obtain an orthogonal transform matrix which is the compact dictionaries of this cluster. In all, we can obtain $\{\boldsymbol{D}_i\}_{i=1}^{K}$.

Then for the i-th patch $\boldsymbol{x}_i$ to be represented, the nearest dictionary $\boldsymbol{D}_{t_i}$ is adaptively assigned to it according to the estimated local orientation, and its coefficient can be readily calculated by $\boldsymbol{\alpha}_i=\boldsymbol{D}_{t_i}^{\mathrm{T}}\boldsymbol{x}_i$.

1.4 Consistency of gradients

To achieve a further improvement, we consider the consistency of the gradients (CG)[22] between the observed LR image $\boldsymbol{Y}$ and the estimated HR image $\boldsymbol{X}$, which can be mathematically defined as

$$\Psi(\nabla\boldsymbol{Y},\boldsymbol{Q}(\nabla\boldsymbol{X}))=\sum_{j=1}^{4}\left\|\nabla_j\boldsymbol{Y}-\boldsymbol{Q}(\nabla_j\boldsymbol{X})\right\|_2^2 \tag{6}$$

The gradients are calculated without considering the blurring operator, since this term is originally designed to preserve more fine edges and produce a sharper output. Gradient operators in four directions are computed according to

$$\begin{aligned}\nabla_1\boldsymbol{X}&=(\boldsymbol{S}_x^1+\boldsymbol{S}_x^{-1}-2\boldsymbol{I})\boldsymbol{X}\\ \nabla_2\boldsymbol{X}&=(\boldsymbol{S}_y^1+\boldsymbol{S}_y^{-1}-2\boldsymbol{I})\boldsymbol{X}\\ \nabla_3\boldsymbol{X}&=(0.5\boldsymbol{S}_x^{-1}\boldsymbol{S}_y^1+0.5\boldsymbol{S}_x^1\boldsymbol{S}_y^{-1}-\boldsymbol{I})\boldsymbol{X}\\ \nabla_4\boldsymbol{X}&=(0.5\boldsymbol{S}_x^1\boldsymbol{S}_y^1+0.5\boldsymbol{S}_x^{-1}\boldsymbol{S}_y^{-1}-\boldsymbol{I})\boldsymbol{X}\end{aligned} \tag{7}$$

where $\boldsymbol{S}_x^{\pm1}$ and $\boldsymbol{S}_y^{\pm1}$ represent the matrix operators of shifting ±1 pixel in the horizontal and the vertical directions, respectively. The calculation of $\nabla_j\boldsymbol{Y}$ is the same as that of $\nabla_j\boldsymbol{X}$.

1.5 Summation and implementation

By incorporating the regularization term of consistency of gradients defined in (6) into the sparse representation-based SR scheme in (5), the final objective function in this paper is

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\alpha}\{\left\|\boldsymbol{Y}-\boldsymbol{QH}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_i\right)^{-1}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_{t_i}\boldsymbol{\alpha}_i\right)\right\|_2^2+\sum_{i=1}^{M}\|\boldsymbol{W}_i\boldsymbol{\alpha}_i\|_1+\eta'\sum_{j=1}^{4}\left\|\nabla_j\boldsymbol{Y}-\boldsymbol{Q}(\nabla_j\boldsymbol{X})\right\|_2^2\} \tag{8}$$

Merging the 1st and the 3rd terms in Eq. (8) together, it can be rewritten as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}}\left\{\left\|\begin{bmatrix}\boldsymbol{Y}\\ \eta\nabla_1\boldsymbol{Y}\\ \vdots\\ \eta\nabla_4\boldsymbol{Y}\end{bmatrix}-\begin{bmatrix}\boldsymbol{QH}\\ \eta\boldsymbol{Q}\nabla_1\\ \vdots\\ \eta\boldsymbol{Q}\nabla_4\end{bmatrix}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_i\right)^{-1}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_{t_i}\boldsymbol{\alpha}_i\right)\right\|_2^2+\sum_{i=1}^{M}\|\boldsymbol{W}_i\boldsymbol{\alpha}_i\|_1\right\} \tag{9}$$

By defining

$$\begin{bmatrix}\boldsymbol{Y}\\ \eta\nabla_1\boldsymbol{Y}\\ \vdots\\ \eta\nabla_4\boldsymbol{Y}\end{bmatrix}=\boldsymbol{Y}',\quad\begin{bmatrix}\boldsymbol{QH}\\ \eta\boldsymbol{Q}\nabla_1\\ \vdots\\ \eta\boldsymbol{Q}\nabla_4\end{bmatrix}=\boldsymbol{H}' \tag{10}$$

Eq. (9) can be changed into a briefer form as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\alpha}\{\left\|\boldsymbol{Y}'-\boldsymbol{H}'\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{R}_i\right)^{-1}\left(\sum_{i=1}^{M}\boldsymbol{R}_i^{\mathrm{T}}\boldsymbol{D}_{t_i}\boldsymbol{\alpha}_i\right)\right\|_2^2+\sum_{i=1}^{M}\|\boldsymbol{W}_i\boldsymbol{\alpha}_i\|_1\} \tag{11}$$

There is no closed-form solution for the above weighted l_1-minimization problem. However, it can be solved effectively by the iterative shrinkage algorithm[23]. The outline of the process for solving (11) is given below.

Step 1 Initialization.

The initial estimate of $\boldsymbol{X}$, denoted by $\hat{\boldsymbol{X}}^{(0)}$, is calculated via applying bicubic interpolation to the input LR image $\boldsymbol{Y}$. Train the geometric dictionaries and determine the initial selection according to subsection 1.4.

Set the initial values: η, P, the maximal iterations L and iteration counter $l=0$.

Step 2 Iterate l until $l>L$ is satisfied.

Compute $\boldsymbol{X}^{\text{tmp}}=\hat{\boldsymbol{X}}^{(l)}+\boldsymbol{H}'^{\mathrm{T}}(\boldsymbol{Y}'-\boldsymbol{H}'\hat{\boldsymbol{X}}^{(l)})$ and $\boldsymbol{\alpha}^{\text{tmp}}=[\boldsymbol{D}_{t_1}^{\mathrm{T}}\boldsymbol{R}_1\boldsymbol{X}^{\text{tmp}},\ldots,\boldsymbol{D}_{t_M}^{\mathrm{T}}\boldsymbol{R}_M\boldsymbol{X}^{\text{tmp}}]$.

Update $\hat{\boldsymbol{\alpha}}^{(l+1)}=S_{\boldsymbol{W}}(\boldsymbol{\alpha}^{\text{tmp}})$, where $S_{\boldsymbol{W}}$ is the soft-thresholding function defined in Ref. [23] and $\boldsymbol{W}$ is the concatenation of the diagonal entries in all $\boldsymbol{W}_i$; and update $\hat{\boldsymbol{X}}^{(l+1)}$ using (4).

In order to save computational cost, the selection of dictionaries is updated every P iterations instead of every time.

2 Experimental Results

To demonstrate the effectiveness and robustness of the proposed single frame SR method in terms of both visual perception and objective criteria, two sets of experiments are conducted. In each of the experiments, our method is compared with three state-of-the-art methods, namely the classic bicubic interpolation method, the example and sparse coding based SR (SCSR)[11], and statistical model based SR (SMSR)[14]. Moreover, to show the effectiveness of CG, the basic method in this paper (without CG) is also added to the comparisons.

2.1 Experimental settings

For dictionary learning, an approximate amount of 80 000 patches of size 7×7 are tailored randomly from several high-quality training images in the first place. Then our geometric dictionaries are trained with $K=183$.

In order to obtain the degraded LR images that are used as input in the following synthetic experiments, six different types of natural images (Butterfly, Lena, Parrot, Parthenon, Pepper and Zebra) are selected, which are commonplace samples in the literature of single frame SR. Then, these images are blurred via a Gaussian filter of size 7×7 with standard deviation 1.6, followed by a down-sampling process with a scaling factor of 3 in both horizontal and vertical directions. The local patches extracted from the HR image are still of size 7×7 with 6-pixel overlaps between adjacent ones. Empirically set $\eta=0.01$, $P=40$ and $L=1\ 400$. For colour images, all of the testing methods are only imposed on the luminance component, since the human visual system is more sensitive to changes in luminance. The chromaticity components are simply interpolated using bicubic interpolation from the input LR image to the target HR image. To evaluate the performance of SR algorithms, the reconstructed images produced by various methods are contrasted in terms of both visual qualities and two numerical indicators (i.e., PNSR and SSIM[24]). For a fair comparison, image borders are included in the calculation of both two numerical indicators, which is omitted in Ref. [14].

2.2 Experiment 1: comparisons on noiseless images

In this part, all the methods are evaluated under the noiseless condition, i.e., suppose that $v=0$. Due to the limitation of content, only one visual comparison on Parrot image is graphically shown in Fig. 1. As we can see from it, bicubic interpolation recovered an excessively smoothed image which is often unacceptable in real applications. The basic method performs a little better than SMSR, while both achieve better results than SCSR on sharpness preserving. By adding the CG term, the proposed method further improves the reconstructed image in the aspect of producing more fine structures and sharper image edges, leading to the most satisfying visual quality. All the PNSR and SSIM results are shown in Tab. 1, from which a consistent conclusion that the proposed method obtains the highest PSNR and SSIM can be obtained.

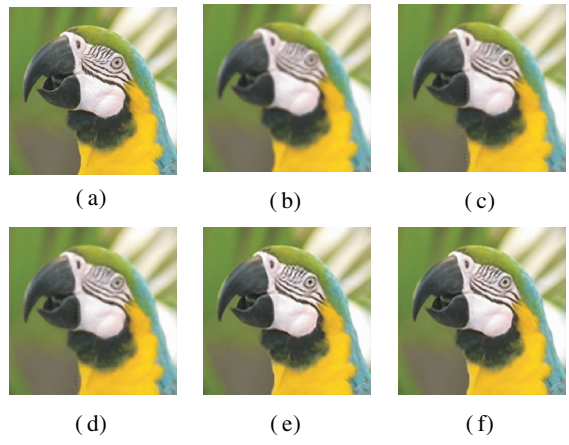

Fig. 1 Visual comparison of noiseless Parrot by different methods with an up-scaling factor of 3. (a) Original image; (b) Bicubic; (c) SCSR; (d) SMSR; (e) Basic method; (f) Proposed method

Tab. 1 Numerical comparisons of PSNR and SSIM results of recovered images under noiseless condition

Method	Butterfly		Lena		Parrot		Parthenon		Peppers		Zebra		Average	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	22.46	0.768 0	28.91	0.808 7	26.76	0.853 3	25.29	0.657 3	29.68	0.848 0	24.79	0.7224	26.32	0.776 3
SCSR	23.90	0.798 3	29.74	0.824 8	28.00	0.868 9	25.85	0.684 0	30.30	0.851 5	26.29	0.768 1	27.35	0.799 3
SMSR	26.51	0.902 2	30.74	0.877 6	28.82	0.910 5	26.64	0.738 0	31.32	0.886 8	28.24	0.842 6	28.71	0.859 6
Basic method	26.72	0.894 7	31.68	0.874 8	29.87	0.909 8	26.79	0.736 8	34.28	0.887 1	28.68	0.846 5	29.67	0.858 3
Proposed method	27.31	0.907 8	31.87	0.878 1	30.21	0.913 0	26.97	0.741 9	34.56	0.889 1	29.06	0.852 3	30.00	0.863 7

2.3 Experiment 2: comparisons on noisy images

Due to the fact that the input LR images are often contaminated by noise in practice, it is necessary to test the robustness of these methods against noise. So in this part, Gaussian white noise with a standard deviation of 5 is further added to the same six input LR images that are used before. Tab. 2 and Fig. 2 show all the numerical results and one of the visual comparisons, respectively. In contrast to the previous performance, the competitive method SMSR results in not only serious noise-caused artifacts in recovered images but also very severe drops in numerical indicators. The reason for this is that a specific set of parameters in SMSR is trained to work efficiently under certain conditions. If the circumstances are slightly changed, the parameters all need to be retrained (reset). In other words, its parameters are so sensitive to the changes of blurring and noise, which is its fatal drawback. On the contrary, the proposed method still leads to quite decent results with a reasonable drop of approximate 1.3 dB on PNSR compared to those in the previous subsection. Note that here we still employ the same settings in Experiment 1 without any change in either the dictionaries or parameters. Moreover, from the aspect of visual sense, the proposed method not only suppresses noise but also preserves more delicate edges in the reconstructed image than any other methods, showing a strong robustness to noise.

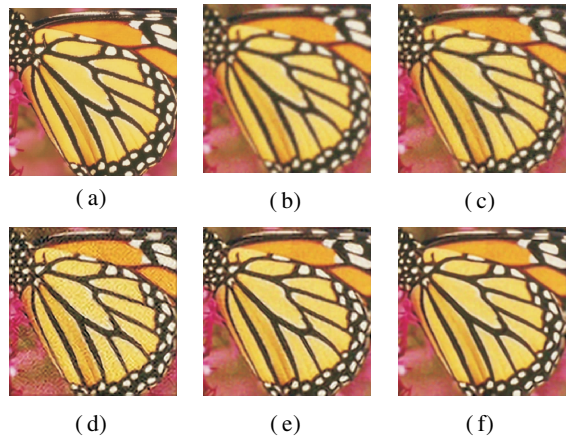

Fig. 2 Visual comparison of noisy butterfly by different methods with an up-scaling factor of 3. (a) Original image; (b) Bicubic; (c) SCSR; (d) SMSR; (e) Basic method; (f) Proposed method

Tab. 2 Numerical comparisons of PSNR and SSIM results of recovered images under noisy conditions

Method	Butterfly		Lena		Parrot		Parthenon		Peppers		Zebra		Average	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	22.40	0.751 1	28.62	0.782 3	26.60	0.823 3	25.17	0.640 2	29.36	0.821 0	24.69	0.709 5	26.14	0.754 6
SCSR	23.74	0.767 4	29.05	0.766 3	27.53	0.797 6	25.56	0.643 7	29.54	0.788 5	25.97	0.736 3	26.90	0.750 0
SMSR	22.94	0.644 2	24.94	0.513 3	24.67	0.507 3	23.47	0.478 8	25.21	0.630 2	24.07	0.753 7	24.22	0.587 9
Basic method	25.61	0.838 8	30.12	0.821 1	28.96	0.874 0	26.11	0.692 7	32.47	0.842 2	27.54	0.785 0	28.47	0.809 0
Proposed method	26.13	0.866 5	30.30	0.824 7	28.96	0.874 4	26.26	0.695 5	32.85	0.851 7	27.82	0.787 2	28.72	0.816 7

3 Conclusion

In this paper, a sparse representation-based single frame super-resolution reconstruction method is proposed. The proposed method utilizes local orientation estimation based image patch clustering combined with the PCA algorithm to learn a series of the geometric dictionaries of different orientations, from which the dictionary of nearest orientation is selected for each local patch to be represented in the sparse coding process. Additionally, the CG regularization term, which can better preserve image edges, is further incorporated into the basic model to obtain a further improvement in the visual quality of reconstructed images. The iterated shrinkage algorithm is adopted to give a mathematical implementation. Extensive experiments on noiseless and noisy images demonstrate that the proposed method can achieve much better results than some state-of-the-art algorithms in terms of both numerical indicators and visual perception.

References

[1] Zeng W L, Lu X B. A generalized DAMRF image modeling for superresolution of license plates[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2012, **13**(2): 828 − 837. DOI: 10.1109/tits.2011.2180714.

[2] Park S, Park M, Kang M. Super-resolution image reconstruction: A technical overview[J]. *IEEE Signal Processing Magazine*, 2003, **20**(3): 21 − 36. DOI: 10.1109/msp.2003.1203207.

[3] Keys R. Cubic convolution interpolation for digital image processing[J]. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 1981, **9**(6): 1153 − 1160. DOI: 10.1109/TASSP.1981.1163711.

[4] Chan T F, Ng M K, Yau A C, et al. Super-resolution image reconstruction using fast inpainting algorithms[J]. *Applied and Computational Harmonic Analysis*, 2007, **23**(1): 3 − 24. DOI: 10.1016/j.acha.2006.09.005.

[5] Rudin L, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. *Physica D*: *Nonlinear Phenomena*, 1992, **60**(/2/3/4): 259-268. DOI: 10.1016/0167-2789(92)90242-f.

[6] Osher S, Burger M, Goldfarb D, et al. An iterative regularization method for total variation-based image restoration[J]. *Multiscale Modeling & Simulation*, 2005, **4**(2): 460 − 489. DOI: 10.1137/040605412.

[7] Ren Z, He C, Zhang Q. Fractional order total variation regularization for image super-resolution[J]. *Signal Processing*, 2013, **93**(9): 2408 - 2421. DOI: 10.1016/j.sigpro.2013.02.015.

[8] Zeng W L, Lu X B, Fei S M. Image super-resolution employing a spatial adaptive prior model[J]. *Neurocomputing*, 2015, **162**: 218 - 233. DOI: 10.1016/j.neucom.2015.03.049.

[9] Candè E J, Wakin M B. An introduction to compressive sampling[J]. *IEEE Signal Processing Magazine*, 2008, **25**(2): 21 - 30. DOI: 10.1109/msp.2007.914731.

[10] Yang J, Wright J, Huang T, et al. Image super-resolution as sparse representation of raw image patches[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage USA, 2008: 1 - 8.

[11] Yang J, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2861 - 2873 DOI: 10.1109/TIP.2010.2050625.

[12] Wang J, Zhu S, Gong Y. Resolution enhancement based on learning the sparse association of image patches[J]. *Pattern Recognition Letters*, 2010, **31**(1): 1 - 10. DOI: 10.1016/j.patrec.2009.09.004.

[13] Yang S, Liu Z, Wang M, et al. Multitask dictionary learning and sparse representation based single-image super-resolution reconstruction[J]. *Neurocomputing*, 2011, **74** (17): 3193 - 3203. DOI: 10.1016/j.neucom.2011.04.014.

[14] Peleg T, Elad M. A statistical prediction model based on sparse representations for single image super-resolution [J]. *IEEE Transactions on Image Processing*, 2014, **23** (6): 2569 - 2582. DOI: 10.1109/TIP.2014.2305844.

[15] Lu X, Yuan H, Yan P, et al. Geometry constrained sparse coding for single image super-resolution [C]// *IEEE Conference on Computer Vision and Pattern Recognition*. Providence, USA, 2012: 1648 - 1655.

[16] Yang S, Wang M, Chen Y, et al. Single-image super-resolution reconstruction via learned geometric dictionaries and clustered sparse coding[J]. *IEEE Transactions on Image Processing*, 2012, **21**(9): 4016 - 4028. DOI: 10.1109/TIP.2012.2201491.

[17] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries [J]. *IEEE Transactions on Image Processing*, 2006, **15**(12): 3736 - 3745.

[18] Candès E J, Wakin M B, Boyd S P. Enhancing sparsity by reweighted l_1 minimization [J]. *Journal of Fourier Analysis and Applications*, 2008, **14**(5/6): 877 - 905. DOI: 10.1007/s00041-008-9045-x.

[19] Candès E J, Romberg J K, Tao T. Stable signal recovery from incomplete and inaccurate measurements[J]. *Communications on Pure and Applied Mathematics*, 2006, **59** (8): 1207 - 1223. DOI: 10.1002/cpa.20124.

[20] Donoho D L. For most large underdetermined systems of linear equations the minimal l_1 norm solution is also the sparsest solution[J]. *Communications on Pure and Applied Mathematics*, 2006, **59**(6): 797 - 829. DOI: 10.1002/cpa.20132.

[21] Feng X G, Milanfar P. Multiscale principal components analysis for image local orientation estimation[C]// *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, USA, 2002, **1**: 478 - 482.

[22] Li X, Hu Y, Gao X, et al. A multi-frame image super-resolution method[J]. *Signal Processing*, 2010, **90**(2): 405 - 414. DOI: 10.1016/j.sigpro.2009.05.028.

[23] Daubechies I, Defrise M, De Mol C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint [J]. *Communications on Pure and Applied Mathematics*, 2004, **57** (11): 1413 - 1457. DOI: 10.1002/cpa.20042.

[24] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, **13** (4): 600 - 612.

基于稀疏表示的单帧超分辨率重建

谢 超[1,2] 路小波[1,2] 曾维理[3]

([1] 东南大学自动化学院, 南京 210096)

([2] 东南大学复杂工程系统测量与控制教育部重点实验室, 南京 210096)

([3] 南京航空航天大学民航学院, 南京 210016)

摘要:为了有效提高重建后的图像质量,提出了一种基于稀疏表示的单帧超分辨率重建方法. 首先,该方法使用一种基于局部方向估计的图像块聚类和主元分析相结合的字典学习方法来获得一系列具有不同方向的几何字典. 然后,给每一个待处理的图像块自动分配一个具有最近方向的字典,并据此进行稀疏编码. 此外,为了在图像锐化和边缘保持方面取得进一步的提高,将梯度一致性加入提出的基本框架. 在自然图像上进行的2组实验表明:提出的方法在视觉和数字指标方面均优于一些先进的同类方法.

关键词:单帧超分辨率重建;稀疏表示;局部方向估计;主元分析;梯度一致性

中图分类号:TP391