

A Novel Image Superresolution Reconstruction Algorithm Based on Sparse Representation

Aili Wang¹, Xinyuan Wang¹, Yuji Iwahori², Yuan Feng¹ and Na Jiang¹

¹ Higher Education Key Lab for Measuring & Control Technology and Instrumentations of Heilongjiang, Harbin, China

² Dept. of Computer Science, Chubu University, Japan
aيلي925@hrbust.edu.cn

Abstract

Superresolution image reconstruction technique uses single or a series of low-resolution images to reconstruct a high resolution image without changing the hardware devices, while improving image quality and the spatial resolution of the image. High resolution means the image with a higher pixel density, can provide more details. In this paper, a novel image superresolution algorithm based on sparse representation is studied. During over-complete dictionary of the training phase, the proposed method improves two aspects including feature extraction and dimension reduction. In the feature extraction process, combining the second derivative with the gradient direction, we construct a new descent direction to improve gradient method. The convergence speed of the new algorithm is faster than the gradient method and can get better results. Then improved two-dimensional Principal Component Analysis (2DPCA) algorithm is used to reduce the dimension, it could eliminate the correlation of the image lines and column. Experiment results show that this method of image reconstruction is better and faster for high resolution image reconstruction.

Keywords: superresolution reconstruction; sparse representation; Overcomplete dictionary; improved 2DPCA

1. Introduction

Image super-resolution technology does not need to change the existing physical device, as long as the use of appropriate techniques of digital signal processing, we would get the high resolution image. It has a large advantage in technology and cost, so it has been applied in many fields, including HDTV, military remote sensing monitoring, public safety, medical imaging and other fields. Compared with multi-frame image reconstruction, single-frame image super-resolution reconstruction technique would be able to estimate the high-resolution images under the same scene, just requiring a low-resolution image of the actual scene. In some applications, for better meeting with the needs of practical application, super-resolution method based on sparse representation is a new research direction of single-frame image reconstruction method.

For the study of image super-resolution method based sparse representation, the dictionary learning is an important step to decide to rebuild the stage results. Zeyde trains the high-resolution image feature which is directly extracted and trained from the dictionary constructing process [1]. At the same time in the process of image reconstruction, he makes the deal with low resolution images which is to be reconstructed through a variety of reconstruction methods, then he uses the method of sparse representation to get the final result. Zhang firstly used the low resolution samples to train a low resolution dictionary, and then according to the degradation process of image, constraints to produce a high-resolution dictionary matching with low-resolution dictionary. This method reduced training time of over-complete dictionary [2]. Yu first

used the low resolution samples to train a low resolution dictionary, and the high resolution dictionary was approximately calculated by the representation of the coefficient training all the high resolution image block and the low resolution image blocks to be reconstructed in the low resolution dictionary, then used the high resolution reconstruction dictionary achieved to reconstruct a high resolution image, and proved the validity of the algorithm through experiment [3]. Yang proposed the approach that used the similarity and redundancy of images of their own internal information, and the high, low resolution image blocks of dictionary couple trained by Pyramid structure and low resolution image to be reconstructed. This kind of method just trained one dictionary, so compared with the method of the high, low resolution image blocks of dictionary couple, it reduced training time [4, 5].

In this paper, we studied a novel image superresolution method based on sparse representation. The feature extraction and dimensionality reduction method have been improved, in the dictionary structure more faster convergence speed of image reconstruction made visual effect and peak signal to noise ratio better than other several methods.

2. Superresolution Reconstruction Algorithm Description

The Image super-resolution reconstruction based on sparse representation algorithm first uses an image training library to get over-complete dictionary of sparse representation [6-7]. The training of over-complete dictionary need to extract the information of high, low resolution image blocks and use K- singular value decomposition algorithm to train the high-resolution image block for obtaining high resolution dictionary of sparse representation., he estimated the desired high-resolution images by sparse representing and reconstructing the high-resolution image trained from the known low-resolution image. The following figure shows the training and reconstructing process of the high-resolution image of super-resolution image reconstruction based on sparse representation algorithm:

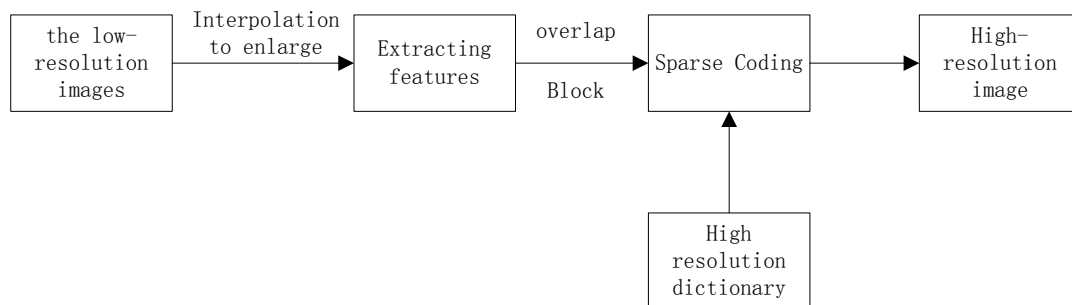


Figure 1. Flowchart for Image Superresolution Reconstruction based on Sparse Representation

The technology of the dictionary training of Super-resolution image reconstruction based on sparse representation algorithm includes K-SVD algorithm and the extraction of training samples. Figure 2 gives the flowchart of dictionary training process.

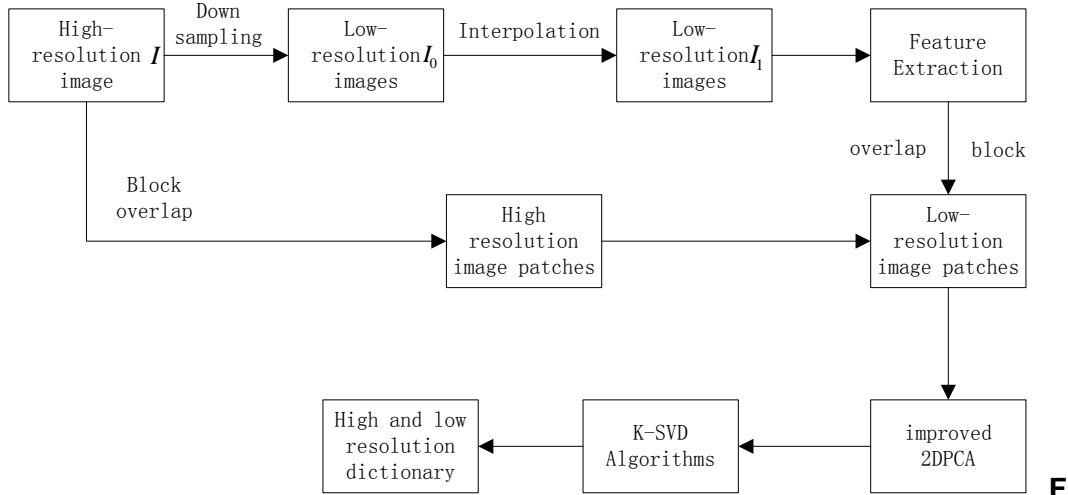


Figure 2. Flowchart of Dictionary Training Process

3. Feature Extraction based on Improved Gradient Method

In the super-resolution image reconstruction algorithms based on sparse representation, image processing is dealt with sub-blocks, the training samples are randomly selected from the image library of high-resolution images^[8]. The number of training samples is usually much larger than the atoms in the dictionary. In order to obtain a wide applicable range of the dictionaries, we should select more natural high-resolution image as an image training database, and the type of the image information of image training database should be as much as possible. Further, we should take into account the dimension of sample in the selection of training samples. In the case of the same number of samples, the larger sample dimension, the longer training dictionary time. This is not conducive to the efficiency of the training dictionary. That cannot very well take into account the large-scale properties of the image. If the dimension of samples is too small, it can not very well take into account the large-scale properties of the image. In this paper, we use sample size in units of blocks in 7×7 for training and reconstruction..

In the training process of specific dictionary, the low-resolution image is degraded from high resolution image. It is often necessary to extract various features from the low-resolution image, divide the features into blocks, and combine them with the corresponding high resolution image to become a vector for training in dictionary. Since the first order and the second order gradient of the image can effectively express the features of the image, and the extraction algorithm is simple, they are often used to represent the features of the image. The first order and second order gradient filter operator of image can be used in the following formula.

$$\begin{aligned} f_1 &= [-1, 0, 1], & f_2 &= f_1^T \\ f_2 &= [-1, 0, -2, 0, 1], & f_3 &= f_2^T \end{aligned}$$

Where T represents transposition. The convolution results of low-resolution image and these four filter operator are regarded as the features of the low resolution image. Then we combine the four features block with the corresponding high resolution image blocks together as the input of the dictionary training algorithm. This method considers the neighborhood information of low resolution image, thus it is benefit to improve the compatibility between reconstructed high-resolution image block.

The second step is performed by a high-pass filter pretreatment to extract small image patches^[9]. The pretreatment of the extracted patch only needs directly employing full image^[10]. This order to solve the problem small patches can avoid the edge effect. As a low-resolution image before treatment, we present a new high-pass filtering method, and the optimization method for gradient algorithm has been improved. When $\frac{\partial^2 f}{\partial x^2} \neq 0$, combine the second derivative with gradient direction, construct a new descent direction

$d = \left[1 + \frac{\delta}{\frac{\partial^2 f}{\partial x^2}} \right] \left(\frac{\partial f}{\partial x} \right)$, where $\delta=1or-1$. This new algorithm converges faster than the gradient method.

4. Improved 2DPCA for Dimensionality Reduction

The advantage of performing a dimensionality reduction is saving computations in the subsequent training and super-resolution algorithms. Therefore, the last step before turning to the dictionary learning stage is reducing the dimension of the input low-resolution patches $\{\tilde{p}_i^k\}_k$. The improved two-dimensional Principal Component Analysis(2DPCA) algorithm is applied on these vectors, seeking a subspace on which these patches could be projected while preserving 99% of their average energy.

The image matrix A size of $m \times n$, $X \in R^{n \times d}$ ($n \geq d$) is a matrix which column vectors are mutually orthogonal, linear transformation $Y = AX$, A is projected onto the image matrix X, the projection Y will have a feature vector of $m \times d$. With total dispersion of the sample as a criterion function projection J(X) can find the best projection matrix X:

$$J(X) = \text{tr}(S_x)$$

Where S_x is the projection matrix Y covariance matrix, $\text{tr}(S_x)$ is the trace of S_x

$$J(X) = \text{tr}\{E[(AX - E(AX))(AX - E(AX))^T]\} \\ = \text{tr}\{X^T E[(A - EA)^T(A - EA)]X\} \quad (1)$$

Image covariance matrix is defined as:

$$G = E[(A - EA)^T(A - EA)] \quad (2)$$

Suppose the number of training samples is M, which is represented as $m \times n$ matrix A_l ($l=1,2,\dots,M$) and the mean of image is:

$$\bar{A} = \frac{1}{M} \sum_{l=1}^M A_l \quad (3)$$

Then G is estimated as follows:

$$G = \frac{1}{M} \sum_{l=1}^M (A_l - \bar{A})^T (A_l - \bar{A}) \quad (4)$$

The largest eigenvalue of the matrix G corresponding to the feature vector X_1, X_2, \dots, X_d constituting the projection matrix X_{opt} X is $X_{opt} = [X_1, \dots, X_d]$ is the optimal value. In the obtained projection matrix $X_{opt} = [X_1, \dots, X_d]$ after the image can feature extraction and classification, and for a given image A.

$$Y_l = AX_l \quad (l=1,2, \dots, d) \quad (5)$$

To obtain a set of feature vectors after projection $L = [Y_1, Y_2, \dots, Y_d]$ it is called the principal component vector image A. In fact, for (3) A_l and \bar{A} , they can be wrote as follow

$$A_l = [(A_l^{(1)})^T (A_l^{(2)})^T \dots (A_l^{(m)})^T]^T \quad (6)$$

$$\bar{A} = [(\bar{A}^{(1)})^T (\bar{A}^{(2)})^T \dots (\bar{A}^{(m)})^T]^T \quad (7)$$

At this time $A_l^{(j)}$, $\bar{A}^{(j)}$ respectively, the first vector A_l and \bar{A} trekking, so

$$G = \frac{1}{M} \sum_{\substack{1 \leq l \leq M \\ 1 \leq i \leq m}} (A_l^{(i)} - \bar{A}^{(i)})^T (A_l^{(i)} - \bar{A}^{(i)}) \quad (8)$$

Now reorder:

$$A_l = [(A_l^{(1)}) (A_l^{(2)}) \dots (A_l^{(m)})] \quad (9)$$

$$\bar{A} = [(\bar{A}^{(1)}) (\bar{A}^{(2)}) \dots (\bar{A}^{(m)})] \quad (10)$$

Here $A_l^{(j)}$, $\bar{A}^{(j)}$ are the jth column and the column vector A_l and \bar{A} so the covariance matrix G based on the image in the column direction can be defined as

$$G = \frac{1}{M} \sum_{\substack{1 \leq l \leq M \\ 1 \leq j \leq m}} (A_l^{(j)} - \bar{A}^{(j)}) (A_l^{(j)} - \bar{A}^{(j)})^T \quad (11)$$

Where $P \in R^{m \times q}$ is a column vector matrix which is perpendicular to each other. A project to the image matrix P, produce a size of matrix $B(q \times n)$

$$B = P^T A \quad (12)$$

The same, it can calculate the former q maximum characteristic value of characteristic vectors $P_1, P_2 \dots P_q$. To get the optimal projection matrix $P_{opt} = [P_1, P_2 \dots P_q]$. So we can use the optimal projection matrix X and P to improve the improved 2DPCA algorithm. The image matrix of A project to the X and P at the same time, it will produce matrix $U(q \times d)$

$$U = P^T A X \quad (13)$$

Matrix U is called characteristic matrix. The training sample image $A_l (l = 1, 2 \dots M)$ projects on the X and P at the same time. to get the training sample characteristic matrix $U_l (l = 1, 2 \dots M)$. Then to select a testing image A to obtain its characteristic matrix U by (13), and then to use the nearest neighbor classifier to classify recognition

5. K-SVD Algorithm Introduction

K-SVD is the core algorithm for sparse representation in the dictionary training process in image super-resolution reconstruction [11]. K-SVD algorithm is also known as the generalized K-means clustering algorithm. When K-SVD algorithm uses only one atom to approximate each signal, it degenerates to K-means clustering algorithm. Its main idea is solving sparse representation of the input sample in the current dictionary and update of the dictionary, alternating these two processes constantly, according to the results of sparse representation of each column to update the dictionary. Suppose there are N training sample as input, over complete dictionary $D \in R^{n \times K}, \{y_i \in R^n, i = 1, 2, 3 \dots N\}$ represents the training sample set, $A = \{\alpha_i \in R^K, i = 1, 2, 3 \dots N\}$ indicates sparse representation coefficients of the training sample set, the K-SVD dictionary training algorithm can be equivalent to solve the equation :

$$\min_{D, A} \{\|Y - DA\|_F^2\} \quad s.t. \quad \forall i, \|\alpha_i\|_0 \leq T_0$$

Where T_0 is the upper limit of non-zero coefficients in sparse representation.

Solving the above equation is an iterative process. First set the initial value of the random dictionary D, and in each iteration, assume the dictionary D is fixed, with the optimization algorithm to obtain input samples sparse representation of the coefficient matrix Y. Finally use the coefficient matrix A to update each column of the dictionary.

If the current dictionary D needs to be updated in the kth column and d_k , in the coefficient matrix A and the dictionary D, in addition to the kth column and d_k of the dictionary and its corresponding coefficient α_k^k representing the outside (that is, the coefficient k line matrix A), the remaining rows and columns are assumed to be fixed, then the multiplicative term in the formula can be rewritten as :

$$\|Y - DA\|_F^2 = \|Y - \sum_{j=1}^K d_j \alpha_j^T\|_F^2 = \|E_k - d_k \alpha_k^T\|_F^2 \quad (14)$$

Where E_k represents the error of N sample with the removal of atoms k.

Updating d_k and α_k^k with SVD, SVD can find the nearest matrix of E_k rank 1, it is not sparse coefficients thus obtained α_k^k . In other words, the location of α_k^k and value of d_k non-zero elements is not the same as before the update. So, intuitively think, leaving only non-zero value coefficients, then SVD decomposition will not appear on this phenomenon. So for E_k and α_k^k do transform, α_k^k nonzero only retain x position E_k and α_k^k retain only those items after the product position and k nonzero. Formation E_k^k will do E_k^k SVD decomposition to update d_k .

6. Experimental Results and Analysis

To demonstrate this algorithm, we use matlab2012b for simulation experiments comparing multiple methods. The patch size used is $n = 81$, and the improved 2DPCA results with a reduction from $4 \times 81 = 324$ dimensions to $n_t = 16$ dimensions. The dictionary training procedure applied 40 iterations of the K-SVD algorithm, with $m = 1000$ atoms in the dictionary, and allocating $L = 3$ atoms for each representation vector.

The test is between our algorithm, bicubic interpolation, yang and ANR methods. The proposed algorithm uses optimized implementation for K-SVD and OMP algorithms.

Around 130,000 training patch-pairs low-resolution dictionary learning takes approximately 10 minutes for 40 iterations Training in the use of high-resolution dictionary for a very short period of time. The results are summarized in the following Table1. Table1 indicates that our proposed algorithm is better than that of bicubic interpolation, methods of yang and ANR and gets better PSNR.

Figure 3 gives the results of baby face image superresolution by different methods in order to evaluate from subject visual aspect. As can be seen from the figure above, from the visual aspect, our algorithm has significantly improved, keeping high-frequency detail much better. Obviously the use of feature extraction combing with second derivative and gradient descent and improved 2DPCA for dimension reduction for trained dictionary in image superresolution reconstruction has played a significant effect, our algorithm is proved to be more optimized.

Table 1. Results of Image Superresolution by Different Methods (PSNR:dB)

name	bicubic	Yang	ANR	Our method
baboon	23.210	23.472	23.623	23.642
barbara	26.249	26.392	26.769	26.937
bridge	24.404	24.824	25.016	25.133
coastguard	26.855	27.015	27.061	27.182
comic	23.116	23.902	23.974	24.026
face	32.820	33.114	33.530	33.647
flowers	27.230	28.249	28.418	29.032
foreman	31.177	32.039	33.198	33.378
Lena	31.678	32.638	32.986	33.128
man	27.008	27.760	27.902	28.035
monarch	29.426	30.713	31.111	31.146
pepper	32.388	33.325	34.046	34.202
ppt3	23.707	24.978	25.232	25.321
zebra	26.634	27.953	28.476	28.667

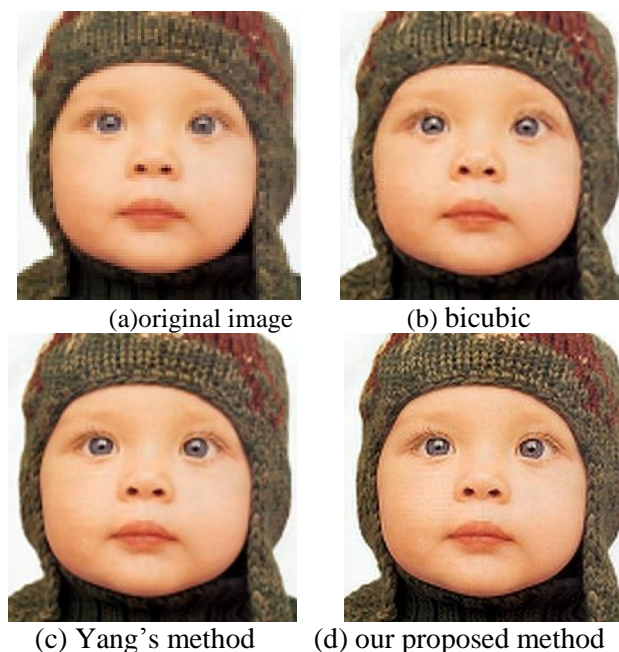


Figure 3. Results of Baby Face Image Superresolution by Different Methods

7. Conclusion

In this paper, we studied the image super-resolution based on sparse representation, mainly improved image feature extraction and dimensionality reduction in the dictionary training process. The use of the second derivative and gradient descent for feature extraction and improved 2DPCA dimensionality reduction, the algorithm for faster dictionary training, and image reconstruction retain more high frequency information.

References

- [1] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches", IEEE Computer Vision and Pattern Recognition (CVPR), June (2008), pp. 1-8.
- [2] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution via sparse representation", to appear in IEEE Trans. on Image Processing, vol. 19, no. 11, (2010), pp. 2861-2873.
- [3] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations", In Curves and Surfaces, (2010), pp. 711-730,
- [4] L. Zhang, M. Yang and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?", In: ICCV, (2011).
- [5] K. Huang, R. Hu and Z. Han, "A Super-Resolution Method Based on Local Sparse and Global Gradient", The third image analysis and signal processing international conference(IEEE-IASP2011), (2011)Wuhan, China, vol. 12, pp. 261-265.
- [6] Y. Zhu, Y. Zhang and H. Li, "An Example-based Method in Multi-frame Super Resolution [C]", Asia-Pacific Signal and Information Processing Association Annual Summit and Conference(APSIPAASC), (2011) Xian, China, vol. 10, pp. 850-855.
- [7] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation, IEEE PAMI, vol. 31, no. 2, (2009), pp. 210-227.
- [8] L. Zhang and X. Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion", IEEE Transactions on Image Processing, vol. 15, no. 8, (2006), pp. 2226-2238.
- [9] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding", IEEE Conference on Computer Vision and Pattern Classification (CVPR), vol. 1, (2004), pp. 275-282.
- [10] J. Wang, S. Zhu, and Y. Gong, "Resolution enhancement based on learning the sparse association of image patches", Pattern Recognition Letters, vol. 31, no.11, January (2010), pp. 234-248.
- [11] R. Rubinstein, M. Zibulevsky and M. Elad, "Efficient Implementation of the K-SVD Algorithm using Batch Orthogonal Matching Pursuit", Technical Report – CS Technique, April (2008), pp. 261-265.

