Learning-based Super-resolution via Canonical Correlation Analysis

Yanzi Wang[1], Jiulun Fan[1], Jian Xu[1, 2], and Xiaomin Wu[1]

1. School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
2. Image Processing and Recognition Center, Xi'an Jiaotong University, Xi'an 710049, China
wyz_zc@126.com, jiulunf@163.com, xujian_paper@126.com, wuxiaomin_0110@163.com

Abstract

The task of image super-resolution is to up sample a low resolution (LR) image while recovering sharp edges and high frequency details. In this paper, a single image super-resolution algorithm via canonical correlation analysis (CCA) is proposed. This method is based on the assumption that the corresponding LR and high resolution (HR) images have high correlation coefficients when transformed into a special space. The proposed approach includes two stages: training and testing. In the training stage, a couple of canonical bases for transformation are calculated with the prepared coupled training sets. In the testing stage, the HR image can be recovered by using the canonical bases obtained in the training stage. In addition, an iterative back projection algorithm is used to further improve the image quality. The experiments demonstrate that our algorithm can reconstruct richer details, with fewer artifacts. Moreover, this algorithm is of less complexity.

Keywords: *Image super-resolution, Canonical Correlation Analysis, Iterative Back Projection, Coherent subspace, Correlation*

1. Introduction

Image super-resolution (SR) attempts to reconstruct a high-resolution (HR) image from one or more low-resolution (LR) images. HR images are needed in many practical applications such as medical image diagnosis, computer vision, satellite imaging and entertainment [1]. Though CCD and CMOS sensors have been widely used to capture HR images, it is hard to obtain an image at a desired resolution level due to the limitations of imaging environments and expensive imaging equipment.

So far, many image SR methods have been proposed. Generally, these methods can be divided into three main classes: interpolation-based SR methods, reconstruction-based SR methods and example-learning-based SR methods. Interpolation-based SR methods determine a function model utilizing the known pixels to calculate the unknown pixels. Some mature methods have already existed, such as linear interpolation and bicubic interpolation [2], etc. Reconstruction-based SR methods are based on prior knowledge of the image degradation to restore HR images. Representative reconstruction-based SR methods include the projection-onto-convex-sets (POCS) approach [3], iterative back projection (IBP) [4], and maximum a posterior (MAP) method [5]. The above interpolation-based and reconstruction-based SR methods are based on a strict and proper mathematical model. Without the limitation of the strict mathematical models, numerous literatures have demonstrated that example-based SR methods are effective. Numerous

literatures have proved that example learning based SR methods can obtain preferable results.

Example-learning-based SR methods, which predict the high-frequency details lost in a given LR image by learning the relationship between the HR and LR images from a training data set. These example-learning-based methods can be divided into three classes: nearest neighbor embedding based SR methods, sparse representation based SR methods and regression based SR methods. Since Freeman *et al.* [6] proposed an example learning-based SR method, which estimated the final HR image patches using the relationship learned by the HR and LR training image patches through Markov Random Field (MRF), numerous example-learning-based SR methods have been reported in the literatures.

Nearest-neighborhood-embedding-based SR methods are based on some assumptions about the space relationship between the LR and HR local patches. Chang *et al.* [7] adopted locally linear embedding (LLE) [8] which assumes the manifolds of small image patches in the LR and HR image with similar local geometry. Under this assumption, the neighbor-embedding-based (NE) method was proposed to estimate the target HR image patches by linearly combining the HR counterparts of neighbors. Numerous scholars have done a lot of improvement on the basis of NE [9-12, 14]. To avoid using external training sets, Glasner *et al.* [10] used the input LR image itself to form training set which starts the research about self-learning. It is based on the observation that the local features of a natural image repeat themselves many times within and across different scales of the original image. Gao *et al.* [11] found the k -nearest neighbors (k -NNs) for linear embedding in a unified feature subspace mapped by LR-HR image patches, which were aimed at reducing the difference between the LR-HR counterparts. Zhang *et al.* [12] partitioned the whole training data set into a set of subsets by clustering the histograms of oriented gradient (HoG) [13], which could effectively reduce computational complexity while preserving SR quality. Recently, Chen *et al.* [14] proposed a method based on low-rank matrix recovery and neighborhood embedding. In Chen's method, the image patches matrix was decomposed into a low-rank component and a sparse component, and then NE algorithm was performed on each component. Obviously, it would inevitably add complexity when the number of training set was increased.

To reduce the complexity, sparse representation-based SR methods, which can represent the images of main information and details using fewer coefficients were developed. These methods consider the corresponding HR and LR images which can be transformed into a unified feature subspace in which these images can be sparsely represented by the same coefficients. Sparse representation was first applied in the super-resolution reconstruction by Yang *et al.* [19]. Mandal *et al.* [20] adopted gradient features to perfect edge information on basis of sparse representation. Zerde *et al.* [21] proposed a method including several important modifications to the local sparse-land model which was similar to Yang *et al.* [19]. Zhang *et al.* [22] learnt a multi-scale dictionary for representing the similar redundancies of local patterns within the same scale and across different scales, which can lead to photo-realistic results.

Another kind of example-learning-based SR method is regression-based SR methods. This kind of method is aimed at building a proper regression model through the training data set. For example, Kim *et al.* [23] adopted kernel ridge regression to learn a map from the LR image to the HR image. Zhang *et al.* [24] employed a non-local-means filter to learn a non-local prior and the steering kernel regression to learn the local prior. Yang *et al.* [25] proposed a fast SR algorithm based on in-place examples, in which a derived function of first-order regression was learnt. In addition, some other approaches have been proposed [26, 27].

At present, CCA has appeared in numerous literatures. The reconstruction stage is operated on the coherent subspace decided by CCA. For example, Huang *et al.* [15] proposed a SR method for the human face. This method performed a NE procedure on a

coherent feature obtained by canonical correlation analysis [18]. Recently, a vehicle logo super-resolution using CCA was reported in An *et al.* [15] to improve the vehicle recognition rate. Instead of one-dimensional CCA, two-dimensional CCA was used to reconstruct a face image [17].

Inspired by the idea of CCA, we will propose learning-based super resolution via Canonical Correlation Analysis (LCCA) in this paper. CCA technology provides a method to map the LR and HR patches into a cosine similarity subspace. In the coherent subspace, the correlation between the LR and HR patch features becomes maximized. We call the mapped LR and HR features as CCA features. In that way, we can reconstruct HR image patches according to the CCA features. In the training phase, we first change the high- and low- resolution training image patches into the transform domain, and obtain the optimal base vectors via canonical correlation analysis. In the testing phase, we reconstruct the test low-resolution image in the transform domain and convert it to the pixel domain to gain the initial result. Finally, an iterative back projection algorithm is used to further improve the image quality. Compared with k -NNs, we get lower computational costs because the proposed method avoids finding nearest-neighbors in millions of training samples. The experimental results show that the proposed LCCA can reconstruct richer details compared with the state-of-the-art SR methods.

The remainder of this paper is organized as follows. In Section 2, we present the proposed method. The algorithm implementation details and the choice of parameters are demonstrated in Section 2. Then, Section 4 shows the quality evaluations and visual effects of the proposed method and compares our method with several other SR methods. Finally, in Section 5 we conclude the paper.

2. The Proposed LCCA algorithm

In this section, we first introduce how to generate the training data, and then we present the proposed method of SR reconstruction in detail.
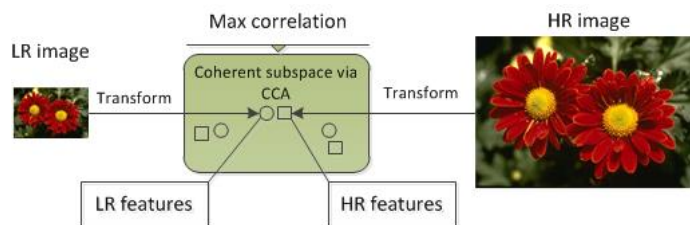

Figure 1. The Schematic of CCA

The main question of restoring an HR image from a LR image is how to establish the relationship between the HR and LR images effectively. To address this problem, we apply Canonical Correlation Analysis to establish the coherent subspace to restore the HR image. Canonical Correlation Analysis (denoted as CCA) was first introduced by Hotelling [18], which was aimed at finding bases for two sets of random vectors such that the correlation between the projections of the vectors onto the bases is maximized. In CCA, the mutual information of two sets is considered. We apply CCA to find the transformation for HR and LR feature spaces that maximize the correlation between the intrinsic geometries of HR and LR images. The schematic of CCA is shown in Figure 1. At present, the sparse representation model has been widely used, and proved very effective. Yang *et al.* [19] first put forward a SR method by coupled dictionaries. Yang's method mapped these feature spaces into the unified feature subspace where they had the biggest similarity based on distance. Unlike the Yang's method, the CCA is used to

determine the coherent subspace, where they have the biggest similarity based on the inner product.

Figure 2 illustrates the flowchart of the whole reconstruction schemes of our algorithm. There are two key steps for single image super-resolution via CCA: training and testing. In the training phase, CCA is used to learn a pair of the best mapping matrices. These matrices are used to reconstruct the HR image from the given test LR image in the testing phase.
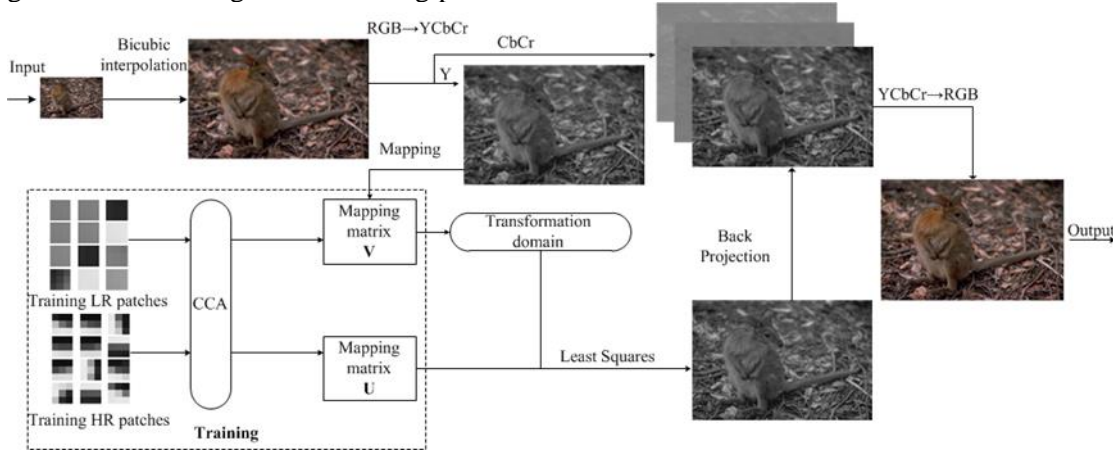

Figure 2. Flowchart of our Proposed Algorithm

2.1. Training Stage

In this subsection, we will introduce the training stage. Among the manifold learning methods, Canonical Correlation Analysis (denoted as CCA) has been widely used in recent years. CCA was first introduced by Hotelling [18], which was a multivariable statistical analysis to describe the relationship between two sets of variables. CCA can find canonical variables of the two sets that have the maximum correlations.

In the training stage, we respectively map the HR and LR image patch feature into a subspace in which the correlation between the canonical variables is maximized. CCA is applied to find the optimal mapping matrices. In our experiments, there are many pairs of HR and LR images in the training set. X_s and Y_s are the corresponding LR and HR features extracted from the training sets, respectively.

We transform the image patch features into column vectors to form the data sets. Hereafter, we obtain a set of training HR and LR image patch feature set: $\{Y_s \in {}^{N_s \times K}, X_s \in {}^{N_s \times P}\}$, where N_s represents the number of LR and HR image patches, K and P represent the dimension of HR and LR image patch feature, respectively. CCA finds a sequence of uncorrelated linear combinations $\{X_s v_m, m = 1, 2, ..., M\}$ for the LR patches $X_s = \{x_s^i, i = 1, 2, ..., N_s\}$, and a corresponding sequence of uncorrelated linear combinations $\{Y_s u_m, m = 1, 2, ..., M\}$ for the HR patches $Y_s = \{y_s^i, i = 1, 2, ..., N_s\}$, such that the correlations in formula (1) are successively maximized.

$$Corr^2(X_s v_m, Y_s u_m) \tag{1}$$

Note that at most $M = \min(K, P)$ directions can be found. Solving the above question, the optimal mapping matrices can be calculated. Then an algorithm is used to calculate the mapping matrices $U = \{u_m, m = 1, 2, ..., M\}$ and $V = \{v_m, m = 1, 2, ..., M\}$.

The CCA solution [28] is computed using generalized Singular Value Decomposition (SVD) of the sample cross-covariance matrix $Y_s^T X_s / N_s$. The optimization problem can be written as:

$$\{u, v\} = \underset{\substack{u^T (Y_s^T Y_s) u = 1 \\ v^T (X_s^T X_s) v = 1}}{\arg \min} \; u^T (Y_s^T X_s) v \tag{2}$$

We can obtain the optimal mapping matrices through Eq. (2). To solve the above optimization problem, first of all, we need to compute the SVD of Q with formula (4) [28],

$$Q = (Y_s^T Y_s)^{-\frac{1}{2}} (Y_s^T X_s)(X_s^T X_s)^{-\frac{1}{2}} \tag{3}$$

The decomposition can be considered as a SVD problem,

$$Q = U^* D V^{*T} \tag{4}$$

where U^* and V^* represent the left and right singular vectors of Q. Secondly, with the learned singular vectors, the solution of Eq. (2) can be given by:

$$
\begin{aligned}
U &= (Y_s^T Y_s)^{-\frac{1}{2}} U^* \\
V &= (X_s^T X_s)^{-\frac{1}{2}} V^*
\end{aligned} \tag{5}
$$

So far, the optimal mapping matrices have been obtained. The procedure of computing the mapping matrices via CCA is summarized in Algorithm 1.

Algorithm 1: Computing the mapping matrices via CCA.

Input:

Training data: Training HR image Y_s and LR image X_s;

Output:

Mapping matrices: U and V

---

1. Up sampling the LR image into an HR version X_s with the bicubic interpolation.

2. Partition Y_s and X_s into a sequence of $q \times q$ image patches with s pixels overlapped in raster-scan order. Construct the training patch normalized feature set $X_s = \{x_s^i, i = 1, 2, ..., N_s\}$ and the intensity set $Y_s = \{y_s^i, i = 1, 2, ..., N_s\}$.

3. Compute the mapping matrices U and V using Eq. (5) .

2.2. Testing Stage

Given a LR test image X_t as an input, we estimate the target HR image Y_t with the help of the computed U and V. The test LR image is represented as a set of small overlapping image patches with the same process in the training stage. For each patch, we construct its feature in the whole image. Then the test LR patch feature set is denoted as $X_t = \{x_t^j, j = 1, 2, ..., N_t\}$, where N_t represents the number of test LR image patches. Mapping X_t onto the transform domain V, we can reconstruct the target HR image intensity set Y_t on the transform domain.

We implement the reconstruction stage on the transform domain. First of all, we map Y_t and X_t onto the mapping matrices U and V, respectively, to obtain the HR and LR image CCA features. The relationship between the HR and LR image CCA features obeys the following:

$$X_t V = Y_t U + E \tag{6}$$

where $E \in \mathbb{R}^{N_t \times K}$ represents the error matrix. To solve the above least square problem, we first write the loss function of residual sum of squares as (assuming E is uncorrelated):

$$RSS(Y_t) = tr((X_t V - Y_t U)^T (X_t V - Y_t U)) \tag{7}$$

Then the derivative of the residual sum of squares is calculated and set to zero. So the least square estimate of the loss function is expressed in the following form:

$$Y_t = X_t V U^T (U U^T)^{-1} \tag{8}$$

To avoid the singular case, we need to add a regularization item, i.e. $U U^T = U U^T + \mu I$,

where μ is set to a small positive value (e.g., 1×10^{-6}), and I is a unit matrix which has the same size with the matrix $U U^T$, noting

$$Y_{t0} = X_t V U^T (U U^T + \mu I)^{-1} \tag{9}$$

Through this procedure, we obtain the HR image patches Y_{t0}.

2.3. Post-processing Procedure

In the previous phase of CCA, as some image patches details may be weakened when they are weighted average in the overlapping pixels, we add the IBP procedure to complement the image details. IBP mentioned in many literatures [11, 12, 14, and 35] can be used to improve the image details. We use the IBP to increase the sharpness and details of the image. With the obtained image Y_{t0}, an IBP [4] algorithm is applied to enhance the quality of the HR result. This method is based on the idea that the recovered image should procedure the same image as the one observed if passing it through the LR image generation model. The generation process of the observed LR image X_t from the original HR image Y_t is

$$X_t = DBY_t \tag{10}$$

where B and D are the operator of blurring and down-sampling, respectively.

So, the final HR image can be obtained from

$$Y_t^* = \arg \min_{Y_t} \|DBY_t - X_t\|_2^2 + \lambda \|Y_t - Y_{t0}\|_2^2 \tag{11}$$

Here, λ is a balancing parameter, whose value (in our experiment, we set it to 1) has little influence on the result, but much on iterative time. The gradient-descent method is utilized to solve the above problem,

$$Y_t^{n+1} = Y_t^n + \eta (B^T D^T (X_t - DBY_t^n) + \lambda (Y_{t0} - Y_t^n)) \tag{12}$$

where Y_t^n denotes the HR image after the n th iteration, and η , whose value (between 0.5 and 1) has little influence on the SR result, λ is the step size of the gradient descent. In our experiments, we set the maximum number of iterations as 20. The procedure of SR reconstruction for a test LR image is summarized in Algorithm 2.

Algorithm 2: SR reconstruction for a test LR image.

Input:

Mapping matrices: U and V ;

Test LR image: X_t ;

Output:

The target HR image: Y_t^*

Up sampling the LR image into a HR version X_t with the bicubic interpolation.

Partition X_t into a sequence of $q \times q$ image patches with s pixels overlapped in raster-scan order.

Calculate each patch's mean value and construct the patch feature set $X_t = \{x_t^j, j = 1, 2, ..., N_t\}$.

Compute the HR image patch intensity set $\{y_t^j, j = 1, 2, ..., N_t\}$ by Eq. (9). Add the mean values to each HR image patch.

Synthesis the HR image Y_{t0} by sequentially merging all HR image patches in $\{y_t^j, j = 1, 2, ..., N_t\}$ with averaging multiple predictions for the overlapping pixels between the adjacent patches;

Enhance the obtained HR image Y_{t0} via IBP algorithm by Eq. (12) to get the target HR image Y_t^*.

3. Experimental Settings and Parameter Selection

In this section, we first introduce the experimental settings, then do experiments on different patch size and determine the best patch size. Thereafter we go to discuss the IBP effect on our proposed algorithm.

3.1 Experimental Settings

In our experiments, we magnify the input LR image by a factor of three. To simulate the real image system, all the training HR images are blurred by a 7×7 Gaussian filter with standard deviation 1 and down-sampled by a factor of 3 and 4 to generate the corresponding training LR images. The training HR images are selected from the software package for Yang *et al.* [19]. Figure 3 shows some sample images from the software package. In our experiments, 100,000 image patches are randomly extracted from the training image pairs to build the training data set. To avoid uninformative image patches affecting the learning efficiency, we exclude the patches with small variances from the training data set. The test images are selected from the Berkeley Segmentation Database [31] and shown in Figure 4.

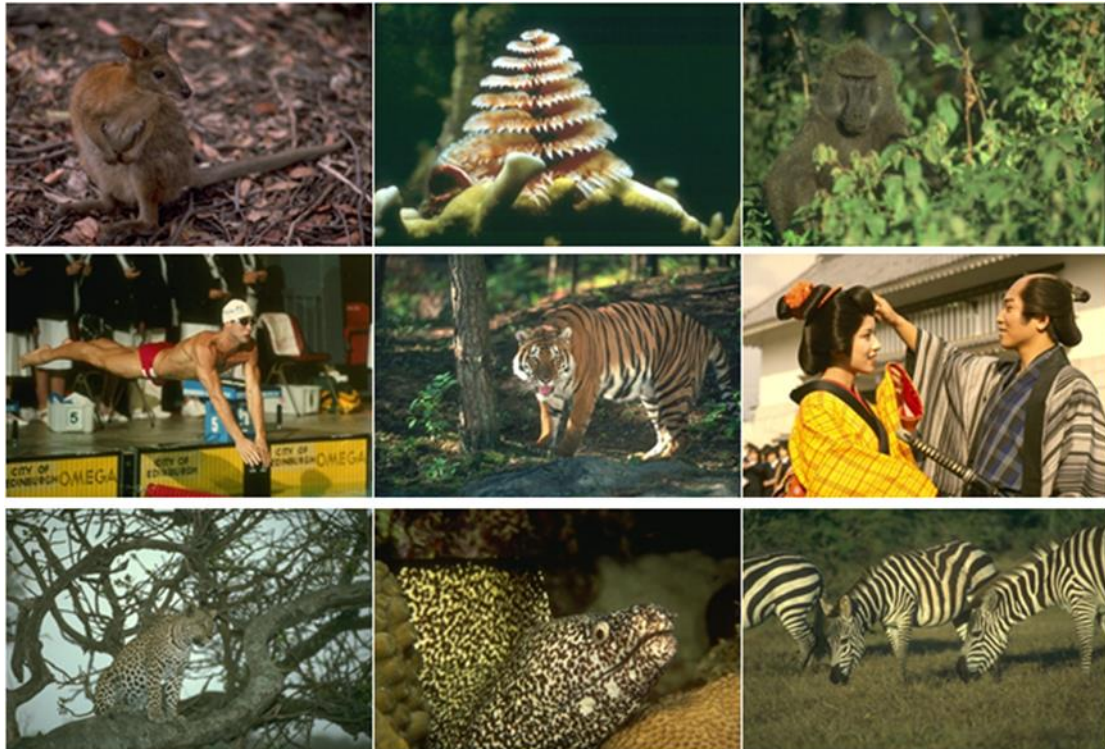**Figure 3. Some Training Sample Images from Dataset**

Figure 4. Test Images; We Refer them from Left to Right and from Top to Bottom as the First to the Ninth Image

For all the training images and testing images, we need to extract their features to learn the best mapping matrices. In some related works, people have suggested extracting different features for the LR image patches. Freeman *et al.* [6] used the high-pass filter to extract the edge information from the LR input patches as the feature. Chang *et al.* [7] and Yang *et al.* [9] used the first-order and second order gradients of the patches as the representation. In this paper, we use the first-order and second-order derivatives as the feature for the LR patches. The four 1-D filters used to extract the derivatives are:

$$f_1 = [-1,0,1], f_2 = f_1^T,$$

$$f_3 = [1,0,-2,0,1], f_4 = f_3^T.$$

(13) where the superscript "T" means transpose. Applying the four filters generates four vectors for each patch, which are concatenated into one vector as the final representation of the LR patch. For the HR image patches, we vector the intensity values of pixels in the HR image patch and subtract its mean value to receive the HR patch feature intensity set.

To validate the effectiveness of the proposed method, we conduct experiments on many sorts of color images. Although Root-Mean-Square Error (RMSE) and Peak Signal to Noise Ratio (PSNR) are widely used as the image quality evaluation indicator, the image quality represented by PSNR-value is not completely consistent with the human intuition. Therefore, we add structural similarity (SSIM) [29] and feature similarity (FSIM) [30, 32] criterion as the image quality evaluation indicators. Since human eyes are more sensitive to luminance component than chrominance components, we transform RGB values to YCbCr color space, and the SR reconstruction is only performed on the luminance(Y) channel. We directly magnify the chrominance (Cb and Cr) channels by the bicubic interpolation.

3.2 Effects of Patch Size

The image patch size could greatly affect the results of example-learning-based SR methods. If the image patch is too big, some details cannot be restored properly. Alternatively, unwanted artifacts will be introduced if too small patch size is used, *e.g.*, noise appears in smooth regions and jagged effects appear on edges in the SR results. In order to avoid the interference of IBP on the parameters, we verify the effects of patch size of the CCA method without IBP. To validate the effect of image patch size, we conduct experiments on different patch sizes of 3×3 with two pixels overlapped, 5×5 with four pixels overlapped, 7×7 with six pixels overlapped and 9×9 with eight pixels overlapped.

To assess the quality of the SR reconstruction objectively, we record the result without the post-processing procedure and utilize four indices, i.e., the RMSE, PSNR SSIM and FSIM to evaluate the performance of the SR results. We compare the results of CCA without IBP with the results of the bicuic interpolation to illustrate the effects of patch size. The curve lines of these four indices are shown in Figure 5. As shown in Figure 5(a), the RMSE-value of CCA-NIBP is bigger than the bicubic when the patch size is 3. The RMSE-values of CCA-NIBP are smaller than the results of the bicubic when the patch size is bigger than 3 and reach the smallest value when the patch size is 7. The inverse tendency could be found in the curve lines for the values of PSNR (Figure 5(b)), SSIM (Figure 5 (c)) and FSIM (Figure 5 (d)).

Figure 6 shows part of the reconstructed SR results of the third test image via different patch size for visual observation. When the patch size is 3, the reconstructed SR image is unsatisfactory. The SR results have lots of artifacts within the image, which is worse than the result of the bicubic interpolation. A lot of details are reconstructed when the patch size is 9×9, but some ghosts are produced. The best reconstructed image is obtained when the patch size is 7×7. Therefore, we use 7×7 in the following experiments.
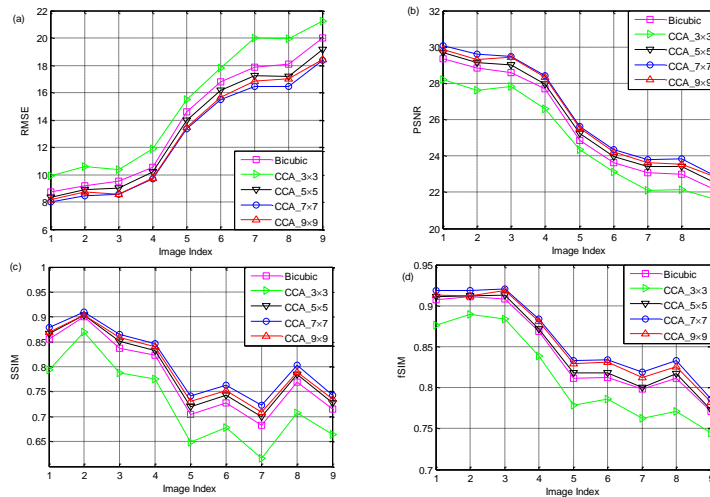

Figure 5. Performance Results of the Nine Images with Different Patch Size for Contrast. (a) RMSE; (b) PSNR; (c) SSIM; (d) FSIM
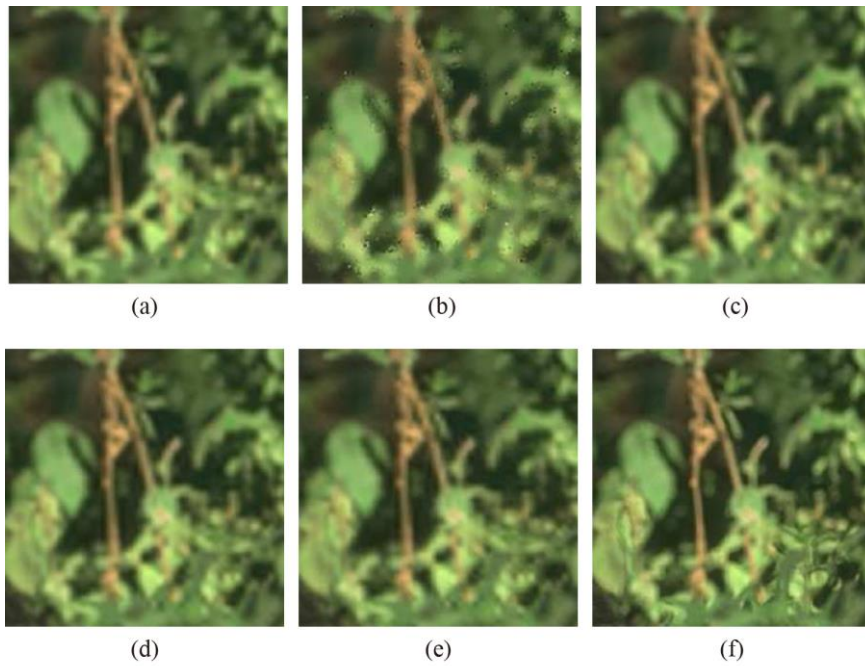
Figure 6. Visual Comparison of Parts of the Third Test Image with Different Patch Size; (a) Bicubic. (b) The Result of 3×3 Patch Size; (c) The Result of 5×5 Patch Size; (d) The Result of 7×7 Patch Size. (e) The Result of 9×9 Patch Size; (f) The Original Image; (Refer to Electric Version and Zoom in for Better Comparison)

3.3 Robustness of CCA

To study the robustness of the CCA procedure, we do experiments on 235 images selected from the Berkeley Database with the CCA procedure, which is performed on the bicubic version of the LR images. We verify the CCA effects compared with the bicubic. As the change of SSIM and FSIM values is little and close to zero, we only utilize two indices, i.e., the RMSE and PSNR to evaluate the effectiveness of the CCA procedure. The statistics histogram of the decrease of RMSE and the increase of PSNR compared with the bicubic results are shown in Figure 7. As shown in Figure 7, for all the 235 images, the results via CCA have a rise in both RMSE and PSNR values. The experimental results indicate that the CCA method can effectively improve the image resolution.
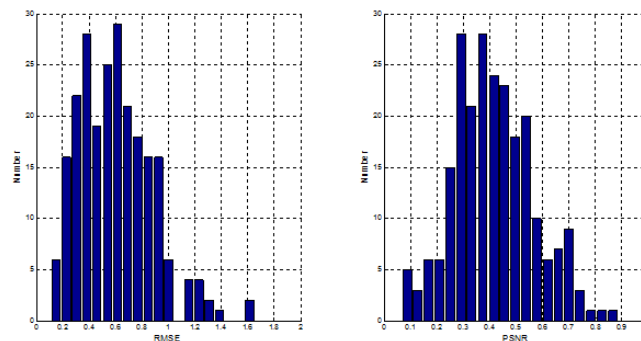

Figure 7. The Decrease of RMSE and the Increase of PSNR Compared with the Bicubic Results

3.4 Effects of IBP

To test the effect of the IBP procedure in the proposed method, we do experiments on the LCCA method before IBP and after IBP. We perform experiments on the nine images shown in Figure 4.Table 1 compares the RMSE, PSNR, SSIM and FSIM values of the proposed algorithm with the bicubic interpolation, CCA without IBP (CCA_NIBP) and CCA with IBP (CCA_IBP). From Table 1, we can see that the IBP procedure has improved the SR performance of our proposed. The IBP method has the ability to improve the image quality, and can restore more details. Visual comparison of part of the third image obtained by CCA_NIBP and CCA_IBP (magnified by a factor of 3) is shown in Figure 8. As show in Figure 8(a), we can observe that the feathers on the head are too smooth, and the nostrils are fuzzy. The image in Figure 8 (b) restores some features on the head but not clear. In addition, we can see richer texture on feathers and much clearer nostrils in Figure 8 (c) than Figure 8(a) and 8(b). Experimental results show that the IBP can enhance details.

Table 1. The Improvement of SR Performance Induced by IBP; For Each SR Method, we have Four Columns; The First Column is RMSE, the Second Column is PSNR, the Third Column is SSIM, and the Fourth Column is FSIM

| Img	CCA_NIBP				CCA_IBP				Increase			
1	8.2494	29.8023	0.8665	0.9108	5.0300	34.0994	0.9324	0.9519	3.2194	4.2971	0.0659	0.0411
2	8.7806	29.2603	0.9049	0.9122	6.3891	32.0221	0.9286	0.9345	2.3915	2.7618	0.0237	0.0223
3	8.8406	29.2012	0.8536	0.9136	4.4382	35.1867	0.9289	0.9612	4.4024	5.9855	0.0753	0.0476
4	10.0122	28.1202	0.8379	0.8757	6.5592	31.7939	0.8989	0.9182	3.4530	3.6737	0.0610	0.0425
5	13.6020	25.4587	0.7282	0.8219	10.3860	27.8019	0.8226	0.8830	3.2160	2.3432	0.0944	0.0611
6	15.8292	24.1416	0.7512	0.8234	12.0415	26.5171	0.8310	0.8719	3.7877	2.3755	0.0798	0.0485
7	17.0892	23.4764	0.7048	0.8051	13.1146	25.7757	0.8084	0.8577	3.9746	2.2993	0.1036	0.0526
8	17.0704	23.4859	0.7866	0.8191	12.9460	25.8881	0.8762	0.8765	4.1244	2.4022	0.0896	0.0574
9	18.6900	22.6986	0.7326	0.7785	13.5058	25.5204	0.8215	0.8384	5.1842	2.8218	0.0889	0.0599
Avg	13.1293	26.1828	0.7963	0.8511	9.3789	29.4006	0.8721	0.8993	3.7504	3.2178	0.0758	0.0481

(a)　　　　(b)　　　　(c)　　　　(d)
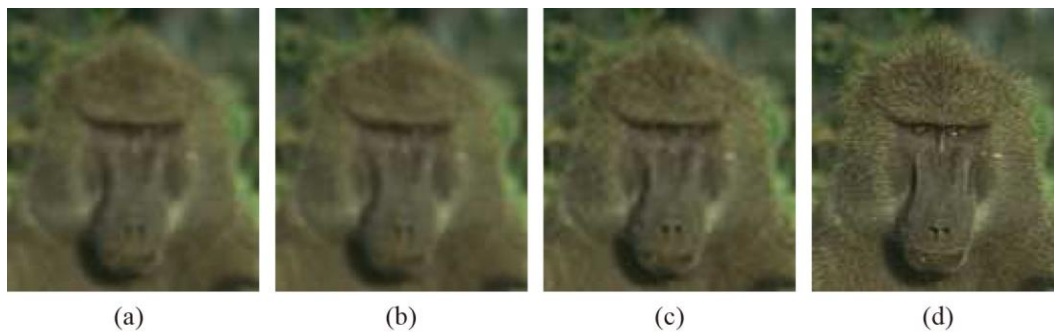
Figure 8. Visual Comparison of 125*141 Portions Extracted from the Third Image Obtained by CCA_NIBP and CCA_IBP (Magnified by a Factor of 3) (a) Bicubic; (b) CCA_NIBP; (c) CCA_IBP; (d) The Original Image; (Refer to Electric Version and Zoom in for Better Comparison)

4. Experimental Results and Analysis

To evaluate the performance of our proposed algorithm, we compare it with the bicubic [2], Yang's method [19] and ANR [34] algorithm, based on the nine test images. First, we will demonstrate of the SR results for factors of 3 and 4. Then, the computational complexity of the proposed algorithm will be discussed.

4.1 3 x Magnification SR Results

In this subsection, we will demonstrate the SR results of the factor of 3. Because the SR process is only done in the luminance channel of the color images, we compare the quantitative difference of this part between the original HR images and the SR outcomes.

We will compare the 3x SR results with bicubic, Yang's method and ANR. Table 2 tabulates the RMSE, PSNR and SSIM values of those methods. As shown, the bicubic gives the worst image quality indices. Yang's method has greatly improved the image quality indices, but not the best. As to the RMSE and PSNR values, the ANR method performs best on four of the nine test images, and the proposed algorithm performs best on the other five images. As to the SSIM values, our proposed algorithm performs best on eight of the nine test images. The quantitative evaluation of the proposed algorithm is superior than others according to the average values.

Table 2. Summary of RMSE, PSNR and SSIM Results of 9 Test Images (3 x Magnifications); For Each SR Method, we have Three Columns; The First Column is RMSE, the Second Column is PSNR, and the Third Column is SSIM

| Img	bicubic			Yang			ANR			CCA		
1	5.8719	32.7552	0.8972	5.4024	33.4791	0.9096	4.9866	34.1747	0.9219	**4.8133**	**34.4819**	**0.9290**
2	7.7336	30.3632	0.9039	6.9827	31.2503	0.9096	6.7060	31.6016	0.9170	**6.6022**	**31.7370**	**0.9204**
3	5.6409	33.1038	0.8943	5.1956	33.8181	0.9025	4.8402	34.4336	0.9145	**4.5980**	**34.8795**	**0.9221**
4	7.9411	30.1332	0.8607	7.0786	31.1319	0.8803	**6.7162**	**31.5884**	**0.8910**	6.8254	31.4483	0.8908
5	12.2368	26.3775	0.7541	11.2307	27.1226	0.7907	11.0703	27.2476	0.8003	**10.9928**	**27.3086**	**0.8086**
6	14.0763	25.1610	0.7682	12.6686	26.0762	0.8103	**12.4818**	**26.2052**	0.8160	12.6631	26.0800	**0.8165**
7	15.4824	24.3340	0.7388	13.7191	25.3843	0.7886	**13.4775**	**25.5386**	0.7958	13.5094	25.5181	**0.7994**
8	16.0440	24.0245	0.8051	14.7731	24.7413	0.8465	14.2151	25.0758	0.8539	**13.4767**	**25.5391**	**0.8687**
9	15.9830	24.0576	0.7580	14.6266	24.8279	0.7848	**14.1845**	**25.0945**	0.8002	14.3858	24.9721	**0.8049**
Av	11.2233	27.8122	0.8200	10.1864	28.6480	0.8470	9.8531	28.9956	0.8567	**9.7630**	**29.1072**	**0.8623**
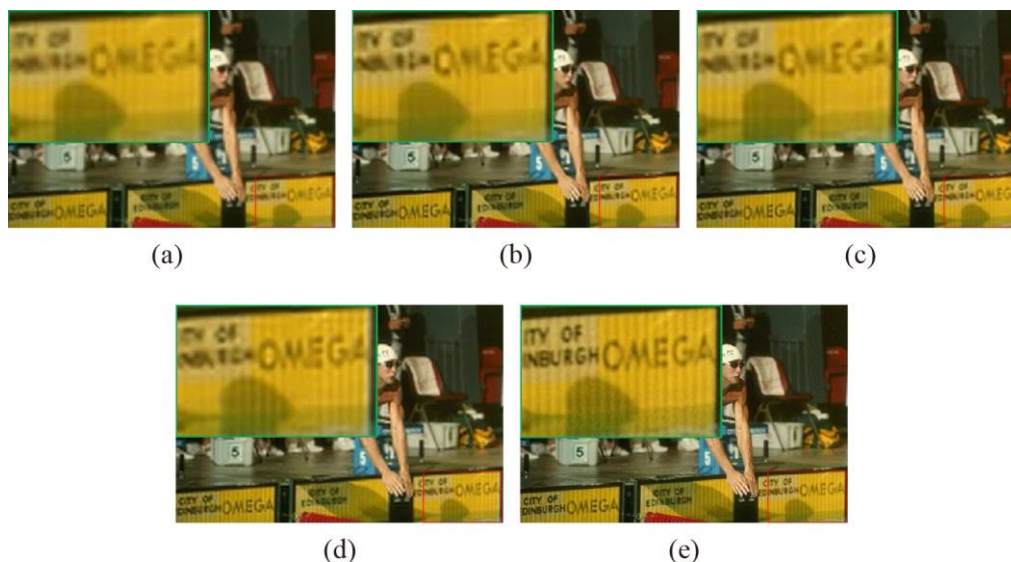
Figure 9. Visual Qualitative Assessment for the Fourth Image with Magnification 3x (a) Bicubic; (b) Yang; (c) ANR; (d) CCA; (e) The Original Image; (Refer to Electric Version and Zoom in for Better Comparison)

Visual examples are shown in Figure 9. As shown in Figure 9, the letters "EGA" from the bicubic result is blurring. The "M" and "A" in Figure 9(b) are blurring that we could hardly see them. The SR result in Figure 9(c) is slightly clearer than the Yang's but the letters are still fuzzy. However, our proposed method can restore clearer letters "MEGA".

Figure 10. Visual Qualitative Assessment for the Flower Image with Magnification 3x (a) Bicubic; (b) Yang; (c) ANR; (d) CCA; (e) The Original Image; (Refer to Electric Version and Zoom in for Better Comparison)

Figure 10 shows the visual comparison on the *flower* image selected from the Berkeley Database. We can observe that the result of the bicubic smooth the stamen's texture. Though Yang's method and the ANR can generate sharp textures, they are not the best. By contrast, the outcomes of our approach can provide richer details.

4.2 4 x Magnification SR Results

In this subsection, we will demonstrate the SR results of 4x magnification. We will compare the 4x SR results with bicubic, Yang's method and ANR. Table 3 tabulates the RMSE, PSNR and SSIM values of these methods. For each SR method, we have three columns. The first column is the RMSE, the second column is PSNR, and the third column is SSIM.

From Table 3, we can see that the proposed CCA and the ANR method perform better than the bicubic and Yang's method. It is similar to the 3x results, as to the RMSE and PSNR values, the ANR performs best on four images of the nine test images, and the CCA performs best on the other five images. As to the SSIM values, the CCA performs best on eight images of the test images. From the average RMSE, PSNR and SSIM values, the proposed method is slightly better than the ANR method.

Figure 11 shows the visual comparison on the tree image selected from the Berkeley Database. Compared with the trunk of the original image, the bicubic and the Yang's results are over smooth, The ANR result reconstructs more textures than other methods, but few than our proposed algorithm. The SR result of our proposed method restores richer details and more robust texture features. Compared with the leaves of the original image, the results of bicubic and the Yang's methods are over smooth. The leaves in the original image are in the same texture direction. The ANR reconstructs some leaves in the vertical direction. From the observations, our proposed method restores sharper edge and richer details in most of the cases.

Table 3. Summary of RMSE, PSNR and SSIM Results of 9 Test Images (4 x Magnifications); For Each SR Method, we have Three Columns; The First Column is RMSE, the Second Column is PSNR, and the Third Column is SSIM

| Img	bicubic			Yang			ANR			CCA		
1	7.7711	30.3211	0.8332	8.1359	29.9227	0.8232	6.7532	31.5406	0.8678	**6.5354**	**31.8254**	**0.8774**
2	9.6150	28.4718	0.8725	9.9494	28.1748	0.8627	**8.5301**	**29.5117**	0.8875	8.5346	29.5071	**0.8889**
3	7.3165	30.8447	0.8371	7.8129	30.2746	0.8282	6.2908	32.1567	0.8675	**6.1947**	**32.2903**	**0.8751**
4	10.1755	27.9797	0.7906	10.4079	27.7835	0.7873	**8.9765**	**29.0687**	0.8203	9.0704	28.9783	**0.8219**
5	14.7898	24.7316	0.6518	14.7292	24.7672	0.6481	13.7245	25.3809	0.6962	**13.5704**	**25.4790**	**0.7081**
6	16.5989	23.7292	0.6805	16.3544	23.8581	0.6897	**15.1214**	**24.5389**	**0.7293**	15.2681	24.4551	0.7266
7	19.2280	22.4521	0.6188	18.9334	22.5862	0.6299	17.5426	23.2489	0.6754	**17.4536**	**23.2931**	**0.6792**
8	20.3141	21.9748	0.6922	20.1677	22.0377	0.6933	18.7041	22.6921	0.7483	**18.1030**	**22.9758**	**0.7656**
9	19.3451	22.3994	0.6709	19.0445	22.5354	0.6651	**17.7245**	**23.1593**	0.7130	17.8399	23.1030	**0.7171**
Avg	13.9060	25.8783	0.7386	13.9484	25.7711	0.7364	12.5964	26.8109	0.7784	**12.5078**	**26.8786**	**0.7844**
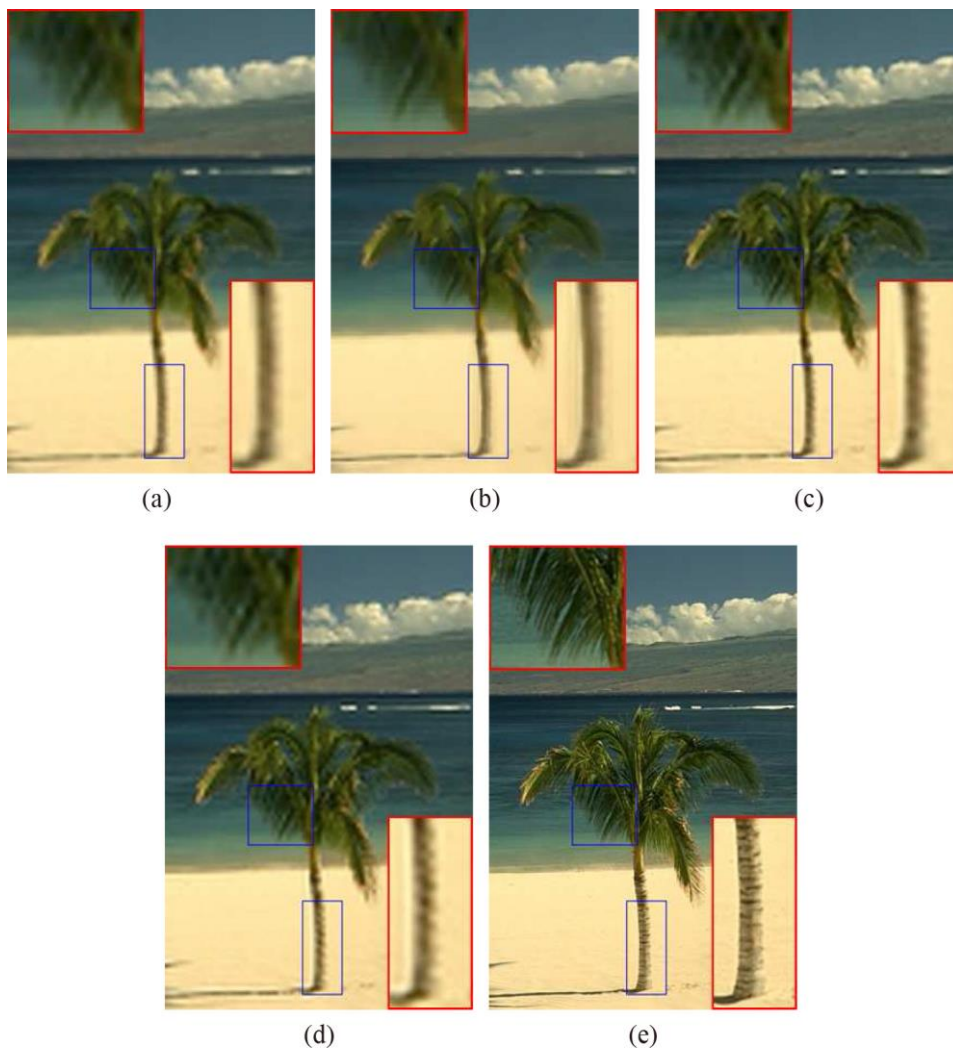
Figure 11. Visual Qualitative Assessment for the Tree Image with Magnification 4x (a) Bicubic; (b) Yang; (c) ANR; (d) CCA. (e) The Original Image; (Refer to Electric Version and Zoom in for Better Comparison)

4.3 Computational Complexity

In this subsection, we will analyze the computational complexity of the training stage and the test stage. As described in Algorithm 1, the training process takes major cost on two parts: computing the mapping matrices U and V. The computation of the mapping

matrices U and V is related to two factors: the number of training image patches N and the super-resolved image patch size s^2 Hence, it totally takes $O(5N^2 s^2) + o(2S^6)$ to obtain U and V. The test process is affected by only one factor: the number of test image patches N_t. It takes $O(3N_t s^4) + O(s^6)$ to reconstruct the target HR image.

We compare the running time with the bicubic, Yang's method and the ANR. We will show the average running time spent on all the test images on an AMD Core8 CPU with 3.10GHz and 8G memory PC at windows platform in Table 4. From Table 4, we can observe that the proposed method takes 64.97s on average, being 2.3 times faster than the Yang's, but 46.7 times slower than the ANR method.

Table 4. The Average Running Times for Different Methods (3x)

Time(s)	bicubic	Yang	ANR	CCA
Avg	0.27	152.57	1.39	64.97

5. Conclusion

This paper presents a CCA-based method to perform single image SR reconstruction. The experimental results of the previous section show that the proposed method balances the performance and the computational complexity. In addition, CCA is a good choice to provide some transformation tools for the machine learning communication which focuses on describing the two related situations (such as illumination change and contrast change). However, one of the most important questions for further study is how to establish more reliable mapping matrices for two related spaces.

Acknowledgements

References

[1] S. C. Park, M. K. Park, M. G. Kang, Super-resolution image reconstruction: a technical overview. IEEE Signal Processing Magzine, vol. 20, no. 3, **(2003),** pp. 21-36.
[2] H. S. Hou and H. C. Andrews, "Cubic splines for image interpolation and digital filtering", IEEE Transaction Signal Processing, vol. 26, 6, **(1978),** pp. 508-517.
[3] A. J. Patti, M. I. Sezan and A. M. Tekalp, "Super-resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time", IEEE Transaction on Image Processing, vol. 6, no. 8, **(1997),** pp. 1064-1076.
[4] M. Irani and S. Peleg, "Improving resolution by image registration", CVGIP: Graphical models and image processing, vol. 53, no. 3, **(1991),** pp. 231-239.
[5] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences", IEEE Transaction on Image Processing, vol. 5, no. 6, **(1996),** pp. 996-1011
[6] W. T. Freeman, T. R. Jones and E. C. Pasztor, "Example-based super-resolution", IEEE Computer Graphics and Applications, vol. 22, no. 2, **(2002),** pp. 56-65.
[7] H. Chang, D.-Y. Yeung and Y. Xiong, "Super-resolution through neighbor embedding", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 275-282 **(2004)** June 27-July 2; Washinqton, DC, USA.
[8] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", Science, vol. 290, no. 5500, **(2000),** pp. 2323-2326.
[9] B. Li, H. Chang, S. Shan and X. Chen, "Locality preserving constraints for super-resolution with neighbor embedding", IEEE 16th International Conference on Image Processing, **(2009)** November 7-10; Cairo, Egypt, pp. 1189-1192.
[10] D. Glasner, S. Bagon and M. Irani, "Super-resolution from a single image", IEEE 12th International Conference on Computer Vision, **(2009)** September 29-October 2; Kyoto, Japan, pp. 349-356.
[11] X. Gao, K. Zhang, D. Tao and X. Li, "Joint learning for single-image super-resolution via a coupled constraint", IEEE Transaction on Image Processing, vol. 21, no. 2, **(2012),** pp. 469-480.

[12] X. Gao, K. Zhang, D. Tao and X. Li, "Image super-resolution with sparse neighbor embedding", IEEE Transaction on Image Processing, vol. 21, no. 7, **(2012)**, pp. 3194-3205.

[13] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, 886-893 **(2005)** June 25-25; San Diego, CA, USA.

[14] X. Chen and C. Qi, "Low-rank neighbor embedding for single image super-resolution", IEEE Signal Processing Letters, vol. 21, no. 1, **(2014),** pp. 79-82.

[15] H. Huang, H. He, X. Fan and J. Zhang, "Super-resolution of human face image using canonical correlation analysis", Pattern Recognition, vol. 43, **(2010),** pp. 2532-2543.

[16] L. An, N. Thakoor and B. Bhanu, "Vehicle logo super-resolution by canonical correlation analysis IEEE 19th International Conference on Image Processing", 2229-2232 **(2012)** September 30-October 3; Orlando, FL.

[17] L. An and B. Bhanu, "Face image super-resolution using 2D CCA", Signal Processing, vol. 103, **(2013),** pp. 184-194.

[18] H. Hotelling, "Relation between two sets of variates", Biometrika, vol. 28(3/4), **(1936),** pp. 312-377.

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

[20] S. Mandal, A. K. Sao, Edge preserving single image super resolution in sparse environment. IEEE 20th International Conference on Image Processing, **(2013)** September 15-18; Melbourne, VIC, pp. 967-971.

[21] R. Zeyde, M. Elad and M. Protter, "On single image scale-up using sparse-representation", Springer Berlin Heidelberg 7th International Conference Curves and Surfaces, vol. 6920, **(2010)** June 24-30; Avignon, France, pp. 711-730.

[22] K. Zhang, X. Gao, D. Tao and X. Li, "Multi-scale dictionary for single image super-resolution", IEEE Conference on Computer Vision and Pattern Recognition, 1114-1121 **(2012)** June 16-21; Providence, RI.

[23] K. I. Kim and Y. Kwon, "Single-image super-resolution using sparse regression and natural image prior", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 32, no. 6, **(2010),** pp. 1127-1133.

[24] K. Zhang, X. Gao, D. Tao and X. Li, "Single image super-resolution with non-local means and steering kernel regression", IEEE Transaction on Image Processing, vol. 21, no. 11, **(2012),** pp. 4544-4556.

[25] J. Yang, Z. Lin and S. Cohen, "Fast image super-resolution based on in-place example regression", IEEE Conference on Computer Vision and Pattern Recognition, 1059-1066 **(2013)** June 23-28; Portland, OR.

[26] R. Liao and Z. Qin, "Image super-resolution using local learnable kernel regression", Springer Berlin Heidelberg 11th Asian Conference on Computer Vision, 7726: 349-360 **(2012)** November 5-9; Daejeon, Korea.

[27] C. Huang, Y. Liang, X. Ding and C. Fang, "Generalized joint kernel regression and adaptive dictionary learning for single-image super-resolution", Signal Processing, vol. 103, **(2014),** pp. 142-154.

[28] J. Friedman, T. Hastie, and R. Tibshirani, "The elements of statistical learning: Linear Methods for Regression", 3rd Ed., Stanford, California **(2001)**.

[29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Transaction Image Processing, vol. 13, no. 4, **(2004),** pp. 600-612.

[30] X. Gao, W. Lu, D. Tao and X. Li, "Image quality assessment based on multiscale geometric analysis", IEEE Transaction on Image Processing, vol. 18, no. 7, **(2009),** pp. 1409-1423.

[31] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics", IEEE 8th International Conference on Computer Vision, vol. 2, **(2001)** July 7-14; Vancouver, BC, pp. 416-423.

[32] L. Zhang, L. Zhang, X. Mou and D. Zhang, "FSIM: A feature similarity index for image quality assessment", IEEE Transaction Image Processing, vol. 20, no. 8, **(2011),** pp. 2378-2386.

[33] R. Fattal, "Image upsampling via imposed edge statistics", ACM Transaction on Graphics, vol. 26, no. 3, article No. 95 **(2007)**.

[34] R. Timofte, V. D. Smet and L. V. Fool, "Anchored neighborhood regression for fast example-based super-resolution", IEEE International Conference on Coputer Vision, 1920-1927 **(2013)** December 1-8; Sydney, NSW.

[35] X. Chen and C. Qi, "Nonlinear neighbor embedding for single image super-resolution via kernel mapping. Signal Processing", vol. 94, pp. 6-22 **(2014)**.