Super-Resolution Image Reconstruction with Improved Sparse Representation

Muhammad Sameer Sheikh and Qunsheng Cao
Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China.
sameer@nuaa.edu.cn

Abstract
In this paper, we present a new approach to reconstruct a high resolution (HR) image from a low resolution (LR) input image based on a two dimensional (2D) sparse method. The new method consists of three phases. Firstly, the nonlinear feature of the input LR image is divided into the linear subspace, and then LR-HR dictionaries are learned to reduce the blurred artifacts of the image. Secondly, 2D sparse representation and self-similarity are developed to strengthen and enhance the image structure. Finally, the final HR image is achieved by reconstruction of all HR patches. Simulation results demonstrated that our proposed method achieved superior results on real images, and shows various improvements in terms of PSNR and SSIM values as compared with some other competent methods.

Keywords: image super-resolution, image enhancement, sparse representation, visual resolution

1. Introduction
Super-resolution (SR) image is considered as a major area of research in digital image processing, and its offer to overcoming the limitation of low resolution images. Many applications have been benefited by this approach such as, biomedical imaging, and astronomical imaging model etc. The aim of super-resolution (SR) image is to reconstruct a visually pleasant high resolution (HR) image from one or more low resolution (LR) input images. In recent year, several SR imaging methods have been proposed, which can be categorized as an interpolation-based method, learning-based and reconstruction-based methods respectively. These techniques provide robust improvement to overcome the low visual resolution in SR imaging [1]. Interpolation-based SR methods are divided into bicubic interpolation and bilinear method respectively. These approach are termed as a simple and fast with less complexity, but they often generate visually displeasing image with sharp edges. The adaptive kernel method is used to evaluate the unidentified pixels present in an HR image grid [2-3], this method are limited to the real-time applications and mostly can generate blurred artifacts at reconstruction part. By using the externally trained datasets, the learning based method is used to reconstruct the HR image by evaluating and mapped between the pairs of LR-HR patches, this method produces blur artifacts and leads to be an unsatisfactory reconstruction because of relying on external training datasets [4].

Reconstruction-based method usually assume the prior knowledge of LR image, which is the combination of several degrading factors, such as blur, noise and down sampling operators. Consequently, many information is missing in a LR image, so one LR image is equivalent to many HR images, this lead to be an ill-posed problem. Recently, different algorithm have been proposed to inverse this problem, by utilizing the prior knowledge, such as, redundancy prior which is used in reconstruction-based method. However, this method can generate visually pleasant image with sharp edges and suppress artifacts [5-6]. Super-resolution image based on 2D sparse has been proposed, this technique is the
robust one and provide satisfactory reconstruction of SR image, but required large processing time and memory [7]. By utilizing the prior knowledge of the image, the problem occurred when recovering the original high resolution image by using less information of LR image [8]. Consequently, this leads to be an ill-posed problem. Yang et. al., [8] have proposed the sparse representation with learning joint dictionary, this approach required large datasets to form the joint dictionaries, and processing of each datasets by $l_1$-norm regularization is still time consuming. Recently, a 2D sparse model has been introduced and applied to enhance the structure of an image [9-10]. Weisheng et. al., [11] proposed the robust method to achieve the HR images by using subdictionary which is constructed from a group of datasets, and select the dictionary adaptively can generate the good resolution but required the large memory and computational time.

Yu et. al., [12] proposed a new method, instead of training datasets the dictionary are learned from the LR input image can suspected to artifacts. Huahua et. al., [13] proposed a new technique based on adaptive co-sparse regularization to improve the SR efficiency by learning sub-dictionary online from the partitioned cluster which can produce good reconstruction but required large amount of memory. Yincheng et. al., [14] proposed a SR image with compressed sensing by using redundancy dictionary to reconstruct the HR image, to generate the redundant dictionary through the training phase need the large amount of calculation, and also this algorithm is limited to static image, so the reconstruction can be unsatisfactory. In this paper, we propose a new method to reconstruct a super-resolution (SR) image from a degraded LR input image. In this framework, the training space is divided into a set of subspace, and then an LR-HR dictionary are constructed. Furthermore, a 2D sparse representation is incorporated to localize and strengthen the image structure. Later, self-similarity is introduced to overcome the repetitive structure and to reduce the artifacts in the image. Finally, the SR image is achieved by reconstructed all HR patches.

2. Problem Formulation and Modeling

![Figure 1. Problem Formulation Block Diagram](image)

2.1. Subspace Modeling

Firstly, we partitioned the features of the LR image into the linear subspace, then obtained the various LR sub-dictionary. We adopt a method to learn a linear relation from a cluster of LR-HR subspace which directly convert the LR feature subspace into the HR subspace [1].

$$Y_l = [y_{l1}, y_{l2}, \ldots, y_{lk}]$$  \hspace{1cm} (1)

$$X_h = [x_{h1}, x_{h2}, \ldots, x_{hl}, \ldots, x_{hk}]$$  \hspace{1cm} (2)

Let suppose the LR feature subspace $Y_l \subseteq \mathbb{R}^m$ and the HR feature subspace is $X_h \subseteq \mathbb{R}^p$. Firstly, construct the datasets from large number of LR-HR image pair. Let the
training samples of LR is \( Y_t = \{ y^1_t, y^2_t, \ldots, y^N_t \} = \{ y^j_i \}_{i=1}^{N_t} \) and the training samples of HR is \( X_h = \{ x^1_h, x^2_h, \ldots, x^N_h \} = \{ x^j_h \}_{i=1}^{N_h} \) respectively, and \( N_t \) represent the number of image pair. Assume \( Y_t^k = \{ y^j_i \}_{i \in \omega_k} \) be the kth subset \( Y_t \), and \( \omega_k \) relate to the index set of \( Y_t^k \). Now we need to divide the HR training set \( X_h \) into K subsets, then \( X_h^k \). The k coupled subsets of LR-HR are formed as \( \{ y^k_i, x^k_h \}_{k=1}^{K} \). Consequently, we can learned the multiple LR-HR dictionary.

### 2.2. LR-HR Dictionary

Multiple LR-HR dictionary are constructed by using the method vectors of LR-HR features which can be represent the linearly in the respective subspaces. Let suppose the given LR image and HR image are \( Y_t, X_h \), respectively, the LR imaging model can be form as,

\[
Y_t = L B X_h + n
\]

where \( L \) and \( B \) are the down sampling operator and blurring operator, respectively. The additive white Gaussian noise is represent as \( n \).

\[
D_l = \arg \min_{D_l} \sum_{i \in \omega_k} \| y^j_i - D^k_l c^k_i \|^2_2
\]

\[
D_h = \arg \min_{D_h} \sum_{i \in \omega_k} \| x^j_h - D^k_h c^k_i \|^2_2
\]

where \( D_l \in \mathbb{R}^{m} \) is the Kth LR sub-dictionary which shows the feature of \( \{ Y_t^k \} \) and \( D_h \in \mathbb{R}^{n} \) is the Kth HR sub-dictionary which shows the feature of \( \{ X_h^k \} \).

Eq. 4 and Eq. 5 are the joint problem, the solution can be achieved by learning the LR sub-dictionary and then transform into the HR sub-dictionary. Where \( c^k_i \) represent the joint coefficient vectors for linear atom in LR and HR dictionary respectively.

\[
c^k_i = (D^k_l D^k_l)^{-1} D^k_l y^j_i
\]

\[
D_l = \arg \min_{D_l} \sum_{i \in \omega_k} \| y^j_i - D^k_l (D^k_l D^k_l)^{-1} D^k_l y^j_i \|^2_2
\]

Similarly,

\[
D_h = \arg \min_{D_h} \sum_{i \in \omega_k} \| x^j_h - D^k_h (D^k_h D^k_h)^{-1} D^k_h x^j_h \|^2_2
\]

Eq. 8 is the optimization problem, in order to achieve the robust solution and to solve the problem of (8). We utilized the LR dictionary feature, the \( y^j_i \) component contained the same features of \( \{ Y_t^k \} \) which can be replaced by defining the Frobenius norm.

\[
D_l = \arg \min_{D_l} \| y^k - D^k_l D^k_l y^k \|_F^2
\]

Eq. 9, represent the LR dictionary, the component \( Y_t \) is a matrix and every column in the matrix is represent the subspace of \( \{ y^j_i \}_{i \in \omega_k} \) and \( \| . \|_F \) is the Frobenius norm. The HR dictionary can be formed as below,
3. 2D Sparse Model

The noisy signal contain the useful information to fully recover the whole signal in the presence of noise is become a critical issue in the digital imaging. Several methods have been proposed to deal with problem such as, sparse representation, this method suspected to generate the artifacts. The 2D sparse method is the robust technique to overcome this problem and to produce the better solution [7].

\[
\sum_{i=1}^{N} \left\| D_i^T B^T D_h^{T} - R x_h \right\|_2^2 + \lambda \| B \|_1
\]

Eq. 12 is the super resolution joint optimization problem with 2D sparse method, first term is the global reconstruction, second and third are sparsity terms which guarantee that every image patch is a sparse representation of both LR and HR dictionaries respectively. \( B^T \) is the sparse matrix, and \( \lambda \) is the coefficient which balance the sparsity term against fidelity, \( \| \cdot \|_2 \).

3.1. Self-similarity Features

Let assume, the analogous patch or pattern exist in the image, for patch \( x_i \) searching for an analogous patch in the image \( x \). Select the patch \( x_i' \), which has the analogous patch of \( x_i \) in the image. The error can be calculated as \( e_i = \| x_i - x_i' \|_2^2 \leq t \), where \( t \) is the threshold [11]. Let assume \( x_i \) is the mid pixel of the patch \( x_i \) and \( x_i' \) is the mid pixel value of the patch \( x_i' \), respectively. Therefore, for predict the patch \( x_i \) by taking the weighted average of \( x_i' \), can be transformed as below,

\[
x_i = \sum_{i=1}^{t} b_i' x_i'
\]

where \( b_i' \) represent the weight which is assigned to the mid pixel \( x_i' \), the updated error shown below,

\[
e_i = \sum_{i=1}^{t} b_i' \| x_i' \|_2^2
\]

Solving (12) for each patch does not guarantee to compatible with adjacent patches and may generate artifacts at the reconstruction part, therefore we are integrating the nonlocal self-similarity features in to (12) to achieve the good resolution patches, and strengthening the image structure with rich detailed.

\[
X = \arg \min_{X,B} \sum_{i=1}^{N} \left\| D_i^T B^T D_h^{T} - R x_h \right\|_2^2 + \lambda \| B \|_1 + z \left\| x_i - \sum_{i=1}^{t} b_i' x_i' \right\|_2^2
\]
3.2. 2D Dictionary

We have learned a 2D dictionaries earlier \( D_t \) and \( D_h \) respectively, the optimization problem can be written as below, ZIM EDIT

\[
B = \arg \min_B \| D_t B^T D_h^T - RX_h \|_2^2 + \lambda \| B \|_1
\] (16)

Eq. 16 is the optimization problem of 2D dictionary, where \( B \) is the sparsing coefficient matrix, and \( X_h \) is the reconstructed HR image patch. Converting in to one 1D dictionary can be written as below,

\[
(D_t, B) = \arg \min_{D_t, B} \| D_t B - X_h \|_2^2 + \lambda \| B \|_1
\] (17)

where \( D_t \) is the joint dictionary that would share the same features of LR and HR image.

4. Image Reconstruction

To achieved an SR image reconstruction. Firstly, scaled up the LR image \( Y_l \) by applying the bicubic interpolation to generate the HR image \( X_h \), which is equivalent to the \( X_i \). Secondly, extract the features of the image by bilateral filter, so it can be defined as \( Y = Y_l R_{kl} \), where \( R_{kl} \) is a bilateral filter which is use to extract the features of \( Y \) from \( Y_l \). Therefore, \( Y \) need to be separated into small patches \( P_k \) as below,

\[
P_k = \sum_{i=1}^{N} R_{ki} Y_{li}
\] (18)

\[
P_h = \sum_{k=1}^{K} P_k = \sum_{k=1}^{K} \sum_{i=1}^{N} R_{ki} Y_{li}
\] (19)

In Eq. 18 the term \( R_{ki} Y_{li} \) can build the HR patches.

\[
X = \arg \min_{X,B} \| P_h - BL X_h \|_2^2 + \sum_{i=1}^{N} \| D_t B - X_h \|_2^2 + \lambda \| B \|_1 + z \cdot \sum_{i=1}^{N} \| x_i - a_i \|_2^2
\] (20)

\[
X_h = X_i + R_{kl}^T R_{kl} P_h
\] (21)

The final image is generated by reconstructed all high resolution patches \( P_h \) and adding \( X_i \) to achieved the HR image.

5. Simulation Results

To evaluate the performance of the proposed method, many experiments have been conducted. In our simulation, we select the different sizes of 60 frames of HR images. In learning 500,000 LR-HR image patches are extracted [7], and then train the 2D dictionary. Therefore, this dictionary represents the LR-HR image patches. We partition the training dataset into 300 clusters. In order to achieve a good reconstruction the selection of patch is very critical issue, we selects the image patch size of 7×7 to reduce the LR features and artifacts.

5.1. Simulation Results with Different Images

We conducts experiments on six different images (cameraman, bike, lena, moon, bag and baboon). Figure 2(a)-(e) are shown the original image, the result of bicubic method, Yang method [8], Yincheng method [14], and the presented method, respectively. The bicubic method often produce low resolution image and generate jaggy artifacts at the reconstructed SR image, which leads to unsatisfactory SR image. Yang [8] method
overcome the limitations of bicubic method by utilizing the overcomplete dictionary, this method can produce blurred artifacts and generate less texture of reconstructed SR image. Since, the processing of overcomplete dictionary consume large time. Yincheng [14] proposed a new method to overcome this problem, and achieved the better SR image by utilizing the training phase which required large calculation to process the redundant dictionary. The proposed method achieved visually pleasing SR image with rich texture detailed and outperformed the competent method.

![Figure 2. Visual Analysis and Comparison of Proposed Method](image)

5.2. Visual Analysis with Different Patch Size

Figure 3(a) - (b) shows the comparison of visual analysis of Lena image with different patch sizes. The sparse representation method and the redundant dictionary method utilized the patch size of (5×5). Since, the small patch has not ability to recover the useful
information of the image [8], [14]. Consequently, this lead to be unsatisfactory SR image with less texture detailed. The presented method achieved robust improvements of reconstructed SR image in terms of visual resolution, PSNR and RMSE values by utilizing the patch size of (7×7).

![Image](image_url)

**Figure 3. Image Analysis with Different Patch Size, (a) 5×5 (b) 7×7**

### 5.3. Simulation Results Comparison in Terms of PSNR and RMSE

Figure 4 is plotted the comparison results of PSNR with different methods by performing on six test images, it can be seen that the proposed method has achieved better average PSNR values and outperforms other competent methods. Also, it has achieved a good texture detailed and produces a good resolution SR image.

Figure 5 is plotted the validation of the proposed method in terms of RMSE. The utilization of the 2D dictionary and good quality patches leads to be a less RMSE values in all of the six images. Furthermore, the proposed method achieved qualitative reconstructed SR image and constantly outperformed other competent methods.

![Graph](graph_url)

**Figure 4. PSNR Comparison of Proposed Method**
6. Conclusions

In this paper, we proposed an image super-resolution (SR) based on the concept of 2D sparse method with learning LR-HR dictionaries. The blurred artifacts of the image were reduced at the reconstruction part by partitioning the LR feature into a cluster form and then, by learning the LR-HR dictionary. The self-similarity features were utilized to enhance the effectiveness of the 2D sparse model, which leads to a good quality of SR reconstructed image. Simulation results conducted on six real images by utilizing the patch size of (7×7), and these results indicated that our proposed method yields a better visual reconstruction of SR image and achieved better PSNR and SSIM values compared with other SR approaches.

References


Authors

Muhammad Sameer Sheikh is currently a PhD candidate at college of Electronics and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China his main research interest focused on the area of computational electromagnetics, radar signal processing, image processing and digital signal processing.

Qunsheng Cao joined the Nanjing University of Aeronautics and Astronautics (NUAA), China, as a Professor of electrical engineering. Dr. Cao’s current research interests are in computational electromagnetics, antenna and microwave technology and the radar signal and image processing. Dr. Cao has published more than 120 academic papers in refereed journals and conference proceedings.