An Effective super-resolution image reconstruction of blurry images using sparse representation method

Atul D. Patil  
Student,  
G.H.R.I.E.M., Jalgaon (MH) India  
Email: atulpatil8286@gmail.com

Swati A. Patil  
Professor and Head,  
G.H.R.I.E.M., Jalgaon (MH) India  
Email: swati.patil@raisoni.net

Abstract

Super-Resolution (SR) imagery plays an important role in many diverse areas of application, such as medical imaging, remote sensing and video surveillance. The super-resolution (SR) is to increase the resolution of an image or a sequence of images beyond the resolving power of the imaging system. Super-resolution is the process of obtaining either a higher resolution still image or a sequence of higher resolution images from a corresponding sequence of low resolution images of a particular scene. In this paper, we deploy the high resolution image from the low resolution image using a powerful statistical image modeling technique. The sparse representation has been successfully used in various image restoration areas as stated above. The sparse representation success is depending on the development of $l_1$-norm optimization techniques and also the natural images are intrinsically sparse in some domain. The resulted high-resolution generated images quality largely depends on whether the employed sparse domain can represent well the underlying image. Extensive experiments on image super-resolution validate that by using adaptive sparse domain selection, the proposed method achieves better results than many state-of-the-art algorithms in terms of both PSNR and structure similarity (SSIM).

Keywords: Sparse representation, image restoration, statistical image, super-resolution, regularization

1. Introduction

Image super-resolution restoration aims to reconstruct a high quality image $x$ from its degraded measurement $y$. Image Restoration is generally modeled as:

$$ y = DHx + v $$

Where $x$ is the unknown image to be estimated, $D$ and $H$ are the degrading operators and $v$ is additive noise. When $H$ and $D$ are identities, the Image Restoration problem becomes denoising, when $D$ is identity and $H$ is a blurring Operator, image restoration becomes deblurring when $D$ is identity and $H$ is a set of random projections, image restoration becomes compressed sensing [1-3]; when $D$ is a down-sampling operator and $H$ is a blurring operator, Image Restoration becomes Super-resolution. The traditional methods of obtaining High Resolution images mainly depend on the hardware aspect and also some time the reducing the pixel size or increasing of the chip size. Actually, increasing the chip size will be followed by an increase in capacitance and reducing the pixel size will lead to an increase in the shot noise. In image restoration, the high cost for high-precision optics and image sensors is also an important concern in many commercial applications regarding finding High Resolution images over the low resolution images. Finally, researchers have explored new techniques to produce high resolution (HR) images from one or multiple frames of low-resolution (LR) imagery, a technique that is called super-resolution (SR) technology. In this paper, we focus on image super-resolution using reconstruction method.

The rest of the paper is organized as follows. Section 2 introduces the related works. Section 3 presents the Super-resolution using sparse representation. Section 4 describes the results. Section 5 concludes the paper.

2. Related Works

As a fundamental problem in image reconstruction and image restoration has been extensively studied in the past three decades [4-10]. It has been found that natural images can be generally coded by structural primitives, e.g., edges and line segments [11], and these primitives are qualitatively similar in form to simple cell receptive fields [12]. In recent years, sparse representation or sparse coding strategy has been widely studied. Partially, due to the progress of $l_0$-norm and $l_1$-norm minimization techniques [13-14].

An important issue of the sparse representation modeling is the choice of dictionary $\Phi$. Much effort has been made in learning a redundant dictionary from a set of example image...
patches. Given a set of training image patches $S = \{S_1, ..., S_N\} \in \mathbb{R}^{N \times N}$, the goal of dictionary learning (DL) is to jointly optimize the dictionary $\Phi$ and the representation coefficient matrix $\Lambda = [\alpha_1, ..., \alpha_N]$ such that $S_i = \Phi \alpha_i$ and $\alpha \leq T$, where $p = 0$ or $1$. To make it tractable the approximation approaches are also including MOD [15] and K-SVD [16].

Regularization has been used in IR for a long time to incorporate the image prior information. The widely used TV regularizations [17] lack flexibilities in characterizing the local image structures and often generate over-smoothed results. In recent years the non-local (NL) methods have led to promising results in various IR tasks, especially in image de-noising and super-resolution. The mathematical framework of NL means filtering was well established.

3. Image Super-resolution using sparse representation

A critical issue in sparse representation modeling is the determination of dictionary $\Phi$. Analytically designed dictionaries, such as DCT, wavelet, curve let and contour lets, share the advantages of fast implementation and they lack the adaptively to image local structures. There are various techniques available such as dictionary learning (DL) methods which provide the learning for universal and over-complete dictionary to represent various image structures. The sparse decomposition over a highly redundant dictionary is potentially unstable and tends to generate visual artifacts. In this paper we propose an adaptive sparse domain selection scheme for sparse representation. In this sparse representation technique the set of compact sub-dictionaries from high quality images are patch in algorithm. These example images patches are clustered into many clusters these clusters consists of many patches with similar pattern and these patterns form the compact sub-dictionary. This compact sub dictionary further can be learned for each cluster. With this sub-dictionary pattern for simplicity we also use the principal component analysis (PCA) technique to learn the sub-dictionaries. Since the given patch can be better represented by the adaptively selected sub-dictionary the whole image can be more accurately reconstructed than using a universal dictionary. For an image patch to be coded the best sub-dictionary that is most relevant to the given patch is selected.

The super-resolution method provides the given low-resolution image $Y$; recover a higher-resolution image $X$ of the same scene. The fundamental constraint with this is that the recovered $X$ image should be consistent with the input $Y$ and $X$ is called Reconstruction constraint. The observed low-resolution image $Y$ is blurred and down-sampled versions of the solution $X$:

$Y = DH X$

Here, $H$ represents a blurring filter, and $D$ the down-sampling operator. The super-resolution technique is extremely ill-posed because for a given low-resolution input $Y$ infinitely many high resolution images $X$ satisfy the above reconstruction constraint. We formulate the problem via the following prior on small patches $x$ of $X$ that is called sparse representation prior and the patches $x$ of the high resolution image $X$ can be represented as a sparse linear combination in a dictionary $D$ of high-resolution patches sampled from training images $x \approx D_k \alpha_k$ for some $\alpha \in \mathbb{R}^k$ with $||\alpha||_0 \ll k$.

To solve the super-resolution problem using the sparse representation prior, the problem is divided into two steps. Initially using the sparse prior we have to find the sparse representation for each local patch respecting spatial compatibility between neighbors. Then using the result from this local sparse representation we have to further regularize and refine the entire image using the reconstruction constraint. In these two step methods, a local model from the sparse prior is used to recover lost high-frequency for local details. The reconstruction constraint for the global model is then applied to remove possible artifacts from the first step and make the image more consistent and natural.

In this section, we propose an Adaptive Sparse Domain Selection scheme, in which a series of compact sub dictionaries are learned and assigns adaptively each local patch a sub dictionary as the sparse domain. With this scheme, a weighted $l_1$-norm sparse representation model will be proposed for IR tasks. Suppose that $\{\Phi_k\}$, $k = 1, 2, ..., K$, is a set of orthonormal sub-dictionaries. Let $X$ be an image vector, and, $X_i = R_i X$, $i =1, 2, ..., N$, be the $i$th patch (size $\sqrt{n} \times \sqrt{n}$) vector of $X$ and $R_i$ is a matrix extracting patch $X_i$ from $X$. For patch $X_j$, suppose that a sub dictionary $\Phi_k$ is selected for it. Then, $X_i$ can be approximated as $X_i = \Phi_k \alpha_i$, $||\alpha_i|| \leq T$, via sparse coding. This type of image $X$ can be reconstructed by averaging all of the reconstructed patches $X_i$ which is mathematically written as:

$\hat{X} = (\sum_{i=1}^{N} R_i^T R_i)^{-1} \sum_{i=1}^{N} (R_i^T \Phi_k \alpha_i)$

In this equation, the matrix to be inverted is a diagonal matrix and hence the calculation can be done in a pixel-by-pixel manner. The image patches are obviously can be overlapped to better suppress noise and block artifacts. We define the...
operator “o” for the convenience of expression using following equation:
\[ \hat{X} = \Phi \alpha + \upsilon \]

Here \( \Phi \) is the concatenation of all sub dictionaries \( \{ \Phi_k \} \) and \( \alpha \) is the concatenation of all \( \alpha_k \).

Let \( Y = DHX + \upsilon \) be the observed degraded image and our goal is to recover the original image \( X \) from \( Y \). Using the ASDS and the definition in given above equation we can easily solve the IR problem which can be formulated as follows:
\[ \hat{\alpha} = \arg \min \{ ||Y - DH\Phi \alpha||_2^2 + \lambda ||\alpha||_1 \}. \]

In above equation it is clear that one key procedure in the proposed ASDS scheme is the determination of \( \Phi_k \) for each local patch. For successfully facilitate the sparse-based Image Reconstruction, we propose to learn offline the sub dictionaries \( \{ \Phi_k \} \), and select online from \( \{ \Phi_k \} \) the best fitted sub dictionary to each patch \( X_i \).

### 3.1 Structure of Sub-dictionaries

In order to learn a series of sub-dictionaries to code the various local image structures, we need to first construct a dataset of local image patches for training [18]. For this purpose, we collected a set of high-quality natural images and cropped from them a rich amount of image patches with size \( \sqrt{n} \times \sqrt{n} \). A cropped image patch denoted by \( S_i \) that will be involved in DL if its intensity variance \( \text{Var}(S_i) \) is greater than a threshold \( \Delta \), i.e., \( \text{Var}(S_i) > \Delta \). This patch selection criterion is to exclude the smooth patches from training and guarantee that only the meaningful patches with a certain amount of edge structures are involved in DL.

Let \( M \) image patches are selected such as \( S = [s_1, s_2, ..., s_M] \). Now our goal is to learn \( K \) compact sub-dictionaries \( \{ \Phi_k \} \) from \( S \) so that for each given local image patch the most suitable sub-dictionary can be selected. Using this local image patch we cluster the dataset \( S \) into \( K \) clusters and learn a sub-dictionary from each of the \( K \) clusters. These \( K \) clusters are expected to represent the \( K \) distinctive patterns in \( S \). If we cropped the hundreds of thousands patches from the training images within that patches many patches are approximately the rotated versions of the others. It allows us to focus on the edges and structures of image patches without taking into account the pixel intensities and helps to increase the accuracy of clustering. Consider, \( S = [s^h_1, s^h_2, ..., s^h_{2h}] \) as the high-pass filtered dataset of \( S \) then we can use the K-means algorithm to partition \( S_h \) into \( K \) clusters \( \{ C_1, C_2, ..., C_K \} \) and denote by \( \mu_k \) the centroid of cluster \( C_k \). Once \( S_h \) is partitioned, dataset \( S \) can then be clustered into \( K \) subsets \( S_k \), \( k = 1, 2, ..., K \), and \( S_i \) is a matrix of dimension \( n \times m \), where \( m \) denotes the number of samples in \( S_k \).

Now we focus on how to learn a sub-dictionary \( \Phi_k \) from the cluster \( S_k \) such that all the elements in \( S_k \) can be faithfully represented by \( \Phi_k \). The design of \( \Phi_k \) can be intuitively formulized by the following objective function:
\[ (\Phi_k, \Lambda_k) = \arg \min \{ || S_k - \Phi_k \Lambda_k ||_F^2 + \lambda || \Lambda_k ||_1 \} \]

Where \( \Lambda_k \) is the representation coefficient matrix of \( S_k \) over \( \Phi_k \) is a joint optimization problem of \( \Phi_k \) and \( \Lambda_k \) and it can be solved by alternatively optimizing \( \Phi_k \) and \( \Lambda_k \) like in the K-SVD algorithm. However, we do not directly use to learn the sub-dictionary \( \Phi_k \) based on the following considerations.

Initially \( l_2 - l_1 \) joint minimization in equation requires much computational cost and then we often assume that the dictionary \( \Phi_k \) is over-completed. Nonetheless, here \( S_k \) is a sub dataset after K-means clustering, which implies that, not only is the number of elements in \( S_k \) limited, but also these elements tend to have similar patterns. We propose to learn a compact dictionary while trying to approximate that equation. So we choose PCA as good solution for this problem. PCA is a classical widely used in pattern recognition and statistical signal processing In this paper, we apply PCA to each sub dataset \( S_k \) to compute the principal components, from which the dictionary \( \Phi_k \) is constructed. Denote by \( \Omega_k \) the co-variance matrix of dataset \( S_k \). By applying PCA to \( \Omega_k \), an orthogonal transformation matrix \( P_k \) can be obtained. It is possible that if we set \( P_k \) as the dictionary and let \( Z_k = P_k^T S_k \), we will then have \( || S_k - P_k Z_k ||_F^2 = || S_k - P_k P_k^T S_k ||_F^2 = 0 \).

### 3.2 Proposed Algorithm

After including both the local AR regularization and the nonlocal similarity regularization into the ASDS-based sparse representation, we have the following ASDS based sparse representation algorithm to solve the IR problem:

#### 3.2.1 Initialization

**a)** Initially, we have to generate the sparse domain by using the wavelet domain then we can compute an initial estimate denoted by \( \bar{X} \) of \( X \) by using the iterated wavelet shrinkage algorithm.

**b)** With the initial estimate \( X_i \), we select the sub-dictionary \( \Phi_k \) and the AR model using \( a_i \) and calculate the NL weight \( b \) for each local patch \( X_i \).
Gaussian noises with standard deviations is kernel used to simulate blurred images. The standard deviation is 3 and the types of blur kernels, uniform blur kernel and different in contents. In the experiments we use two see that these two sets of training images are very quality images often called noisy images. We can high quality images that we called Noiseless results for the luminance component.

4. Experimental Results

To illustrate the robustness of the proposed method to the training dataset, we use two different types of images dataset in the experiments. In the first dataset each image having high quality images that we called Noiseless images. In second type of dataset having low quality images often called noisy images. We can see that these two sets of training images are very different in contents. In the experiments we use two types of blur kernels, uniform blur kernel and Gaussian kernel. The Gaussian kernel having standard deviation is 3 and the 9 × 9 uniform kernel used to simulate blurred images. The Gaussian noises with standard deviations is \( \sqrt{2} \) was then added to the blurred images. For all the test methods we use were applied the color images to the luminance component only because human visual system is more sensitive to luminance changes, and the chromatic components was applied by the bi-cubic interpolator. Finally we generate the report in the form of PSNR and SSIM results for the luminance component.

4.1 Simulated Experiments:

Table 1: PSNR and SSIM results of Different Reconstruction Methods in Simulated Experiments

In the simulated experiments shows in Table 1, the three original HR images used such as “aerial” image, the “spot 5” image and the “cameraman” image with the variable size respectively. Among the three SR algorithms [19], it is shown that the proposed method produces a better SR result. The good performance of the proposed SWTV method can also be illustrated by the PSNR and SSIM values presented in Table 1. It is shown that the proposed method produces the highest PSNR value and also has the highest SSIM value.

4.2 Real Data Experiments

In the real data experiments, we showed the results of three real data sets which are used to verify the proposed algorithm. The reconstruction results of these real data sets are shown in Figure.

**Figure 1:** Reconstruction results of “Leena” image.
(a) Original HR image. (b) LR image. (c) Proposed Result and having uniform blur kernel and \( \sigma_n = \sqrt{2} \). LR image (PSNR = 22.3722) Resulted image (PSNR = 31.1637, SSIM = 0.8848)

**Figure 2:** Reconstruction results of “Parrot” image.
(a) Original HR image. (b) LR image. (c) Proposed Result and having uniform blur kernel and \( \sigma_n = \sqrt{2} \). LR image (PSNR = 23.8711) Resulted image (PSNR = 1.2037, SSIM = 0.8984)
5. Conclusion

Image restoration and reconstruction (IR) is a fundamental topic in image processing and computer vision applications, and it has been widely studied. In this paper, we reconstruct the image using a sparse coding technique. Since the optimal sparse domains of natural images can vary significantly across different images and different image patches in a single image. Improve the quality of reconstructed images autoregressive (AR) models were learned from the training dataset and were used to regularize the image local smoothness. The experimental results on natural images showed that the proposed approach can significantly out perform other leading image reconstruction methods in both parameters such as PSNR and visual quality (SSIM).

6. References


