

Sparse Representation Based Super-Resolution Algorithm using Wavelet Domain Interpolation and Nonlocal Means

G. Suryanarayana*, Ravindra Dhuli

School of Electronics Engineering, VIT University, Vellore-632014, Tamilnadu, India

*Corresponding author, e-mail: surya.gunnam@yahoo.co.in

Abstract

In this correspondence, we propose a novel image resolution enhancement algorithm based on discrete wavelet transform (DWT), stationary wavelet transform (SWT) and sparse signal recovery of the input image. The nonlocal means filter is employed in the preliminary denoising stage of the proposed method. The denoised input low resolution (LR) image is then decomposed into different frequency subbands by employing DWT and SWT simultaneously. In parallel, the denoised LR image is subjected to a sparse signal representation based interpolation. All the estimated high frequency subbands as well as the sparse interpolated LR image are fused to generate a high resolution (HR) image by using inverse discrete wavelet transform (IDWT). Experimental results on various test images show the superiority of our method over the conventional and state-of-the-art single image super-resolution (SR) techniques in achieving the real time performance.

Keywords: *image super-resolution, sparse signal representation, nonlocal means filter, discrete wavelet transform, stationary wavelet transform*

Copyright © 2015 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

High resolution (HR) images are always essential in various digital imaging applications, such as satellite imaging, infrared imaging, medical imaging and text image restoration. In general, the resolution can be increased by increasing the number of pixels per unit area. But the light incident on each pixel reduces there by causing shot noise. Similarly, increasing the chip size also improves the resolution, but results in poor charge transfer rate. So, we rely on effective signal processing techniques for generating a HR image from a given input low resolution (LR) image(s). These techniques are commonly referred to as image super-resolution (SR) algorithms [1-14].

Based on the number of input LR images employed, SR algorithms are categorized into two classes: single image SR and multiple image SR. In multi-frame SR the input LR images should be available with subpixel motion shifts between them. Otherwise, the frames would be coincident causing information redundancy. Hence single image SR has become widely popular in the recent times. These techniques are further classified into three types which are based on interpolation [2-4], machine learning techniques [5-9] and wavelet techniques [10-13].

The well-known interpolation methods like nearest neighbor, bilinear and bicubic are widely used because of their simplicity. But, these techniques suffer from blurred high frequency details due to the ringing and zippering artifacts. Xin and Michael [2] introduced a non-iterative adaptive interpolation scheme based on the local covariance estimates of the input LR image. The algorithm uses the notion of geometric regularity to approximate the relation between the HR and LR covariances. Zhang and Wu [3] suggested interpolating the input LR image in two statistically independent directions. The final reconstructed image is produced by adaptively fusing these two individual results using directional filtering with minimum mean square error. The authors in [4] developed an edge guided image interpolation algorithm based on soft-decision estimation by proposing an auto regressive model. This approach preserves the spatial information to some extent, but the block wise computation of pixels still introduces artifacts. In spite of low computational complexity, the performance of the above interpolation based

approaches degrade quickly when the enhancement factor is higher than 2. In order to achieve SR at large scaling factors, the researchers have focused on single image SR techniques.

The machine learning techniques employ a learning step between numerous HR image patches and its LR counterparts. The learned knowledge is then incorporated into a priori term in the testing phase for SR reconstruction. These priors extract the high frequency details from the training examples holding the known HR components. Machine learning techniques yield good results even at large magnification factors (less than or equal to 3). This improvement is mainly due to the application of dictionary learning and sparse representation in image SR algorithms. Stephane and Guoshen [5] developed a class of linear estimators, which are obtained by adaptively combining a group of estimators derived from different priors. The weights associated with these priors are computed over a block of coefficients by imparting sparse representation. Jiancho et al. [6] proposed a sparse representation model which makes an assumption that each LR patch can be well represented as a sparse linear combination of the columns of the LR dictionary. The same sparse vector is enforced on the HR dictionary in order to reconstruct the HR patch of the same scene. Zeyde et al. [7] improved Jiancho's method with some modifications like, implementing different approach for learning a dictionary pair and boot-strapping scheme for operating without a training database. A more recent work by Dong et al. [9] further contributed to machine learning based SR by proposing a nonlocal centralized sparse representation model. Although most of these techniques perform better, huge number of irrelevant training examples increases the computational time and also deviates the searching process [14].

On the contrary, the wavelet based SR reconstruction techniques provide a better tradeoff between computational complexity and performance. Besides, these SR techniques are able to operate at a magnification factor of 4 or even more, by making use of self-similarities between the local neighborhood regions of the subband images. A notable work done by Demirel and Gholamreza [10] focused on the edge preserving SR by employing discrete wavelet transform (DWT) and stationary wavelet transform (SWT). The immediate substitution of the LR image in place of the low frequency subband leads to non-uniform illumination in the reconstructed HR image, obtained by using inverse discrete wavelet transform (IDWT). Chavez and Ponomaryov [11] suggested an image resolution enhancement technique which generates sharper output image using DWT and sparse mixing estimators. More recently, DWT [13], lifting wavelet transform [12, 15] and dual-tree complex wavelet transform have also been employed in the context of SR.

In our proposed method we have not used the redundant edge preservation stages rather used a stationary wavelet scheme. It minimizes the information loss caused by the inferior directionality and translation varying nature of DWT. Another modification is in applying sparse representation based interpolation on the input LR image to produce the estimated low frequency subband. To preserve the edge information, we developed a normalized edge extraction stage using the high frequency subband images. Finally, we have applied back projection algorithm in order to minimize the SR reconstruction error between the simulated and observed images. The qualitative and quantitative analyses prove the effectiveness of our method over the conventional and state-of-the-art methods.

The remainder of this paper is organized as follows. In section 2 we describe the sparse representation based signal recovery process, the nonlocal means denoising filter and the iterative back projection algorithm. Section 3 illustrates the proposed SR technique followed by various experimental results in section 4. Finally, the paper concludes in section 5.

2. Preliminaries

In this section, we review the preliminary concepts required to develop our algorithm.

2.1. Signal Recovery Based on Sparse Representation

We employ, the sparse representation based interpolation [11] to reconstruct the estimated low frequency subband. The problem is: given a LR image X and we need to recover a HR image Y . It is generally an ill-posed task, since there are infinitely many HR solutions Y or a given LR image. So, the problem is further regularized by enforcing a sparse representation vector, which is usually referred as prior term.

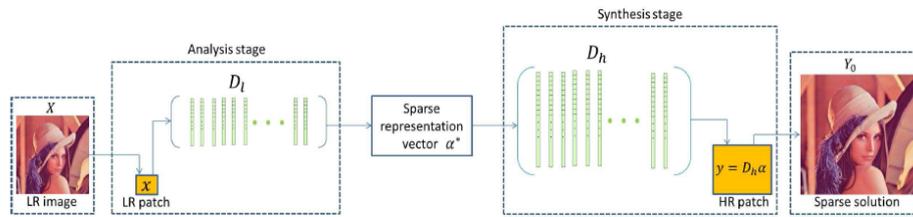


Figure 1. 2-D signal recovery based on sparse representation

Let $D \in R^{m \times k}$ represents an over-complete dictionary having k elements ($k > m$) in m -dimensional space and a signal $y \in R^m$ can be well represented as a sparse linear combination of the columns of D . So, the signal y can be modeled as, $y = D\alpha$, where $\alpha \in R^k$ ($\|\alpha\|_0 \ll m$) is a sparse representation vector. In practice, the LR image can be viewed as a blurred and down-sampled version of the original HR image. Thus, the LR patch x and HR patch y can be mathematically related as $x = My = MD\alpha$, where $M \in R^{k \times m}$ ($k < m$) is a projection matrix that provides control over the blurring and decimation effects.

Since D is an over-complete dictionary, the equations $y = D\alpha$ and $x = MD\alpha$ are underdetermined for the unknown sparse representation vector, α . In spite of that, the sparsest solution α^* will be unique only under mild conditions. In our work, we use the LR and HR dictionaries D_l and D_h advocated in Jiancho's method [6]. These dictionaries satisfy a near-isometry condition, which allows the same sparse representation for an LR-HR patch pair with respect to the D_l and D_h respectively. This process is outlined as a block diagram in Figure 1.

Algorithm for Sparse Recovery

1. **Input:** Single LR image X , two trained dictionaries D_l and D_h .
2. **Process** each 3×3 patch x of X , taken in a raster scan order with one pixel overlap and solve the following optimization problem for α

$$\alpha^* = \arg \min_x \lambda \|\alpha\|_1 + \frac{1}{2} \|D_l \alpha - y\|_2^2 \quad (1)$$

3. **Compute** the HR patch, $y = D_h \alpha^*$ and put into local image Y_0 .
4. **End**
5. **Output:** Sparse recovered solution Y_0 .

This sparse recovered signal Y_0 constitutes the estimated low frequency subband which is used to reconstruct the super-resolved image, by applying IDWT operation.

2.2. Nonlocal Mean Denoising

Prior to the subband decomposition, the input LR image is subjected to the nonlocal mean denoising. The algorithm denoises the LR image based on the assumption that there exist redundant image pixels with in certain neighborhood. Thus it estimates each denoised pixel by computing the weighted sum of the neighborhood pixels in the input image X . The denoised image at location, (p, q) can be viewed as:

$$\hat{X}(p, q) = \frac{\sum_{(i,j) \in Q(p,q)} X[i, j] W[i, j; p, q]}{\sum_{(i,j) \in Q(p,q)} W[i, j; p, q]}, \quad (2)$$

Where $Q(p, q)$ represents the neighborhood of the (p, q) -th pixel and $W[i, j; p, q]$ is the weight associated with it. These weights are computed by taking into account of both the geometric and radiometric distances as:

$$W[i, j; p, q] = \exp \left\{ -\frac{(X[p, q] - X[i, j])^2}{2\sigma_R^2} \right\} \cdot f(\sqrt{(p-i)^2 + (q-j)^2}), \quad (3)$$

where σ_R stands for the range weight and $f(\cdot)$ denotes the geometric distance function. Here the weights are inversely proportional to the radiometric distances. The denoised LR image is further processed in spatial and wavelet domains to compute the super-resolved image.

2.3. Back-projection Algorithm

Any LR image X can be viewed as blurred and decimated version of the ground truth HR image Y . This can be written as $X = HBY$, where H is the down sampling matrix and B is the blurring operator. To minimize the SR reconstruction error, we apply iterative back projection algorithm in the final stage of our proposed method. The algorithm estimates the final HR image by back projecting the obtained SR image Y^* of our method on to the solution space of $X = HBY$, iteratively.

$$Y_{t+1} = Y_t + \mu[B^T H^T(X - HBY_t) + c(Y_t - Y^*)] \quad (4)$$

Here Y_t represents the back projected image at t -th iteration, μ denotes the step size and c is any positive constant. The back projection algorithm is simple and easy to implement and also capable of handling several observations with different degradation parameters. However, it has limitations in terms of the ill posed nature of the SR problem due to absence of unique solution and also its accuracy depends on initial estimation.

3. Proposed method for SR

Let X be the given input LR image of dimensions $r \times s$, using which we have to reconstruct the HR image Y of dimensions $\beta r \times \beta s$. In wavelet based SR techniques, the super-resolved images suffer from non-uniform illumination, which is due to the direct replacement of the low frequency subband with the input LR image. So, preserving the contrast along with the edges is essential. In this correspondence, sparse representation based signal recovery is employed.

Here we start with the input LR image is of dimensions 128×128 . To overcome the inherent noise in the input LR image, we have applied the nonlocal means denoising algorithm. The working of this nonlocal mean filter has been illustrated in section 2. Further, the denoised LR image is subjected to DWT and SWT operations simultaneously, which decomposes the image into four different subbands, viz., low-low (LL), low-high (LH), high-low (HL) and high-high (HH). Due to the down sampling effect, the DWT subbands are decimated by a factor of 2, whereas the SWT subbands have equal dimensions as that of the denoised LR image [10]. The high frequency subbands LH, HL and HH provide information regarding the horizontal, vertical and diagonal components of the image, while the low frequency subband LL being the approximate coefficient. Next we have applied Lanczos interpolation with an up scaling factor of 2 on the DWT high frequency subbands. The down sampling effect of the DWT subbands result in loss of information. In order to minimize this loss, SWT is also employed on the denoised input LR image. Now the interpolated DWT high frequency subband images and the SWT high frequency subbands are added up with each other. The resultant high frequency subbands contain significant frequency components.

To preserve more edge details in the super-resolved image, we developed a normalized edge extraction stage by using the high frequency subbands LH, HL and HH using nearest neighbor interpolation process. The extracted edges can be computed as follows:

$$E = \frac{1}{\gamma} \sqrt{LH^2 + HL^2 + HH^2}, \quad (5)$$

Where γ is the normalization parameter which controls the insignificant edge information.

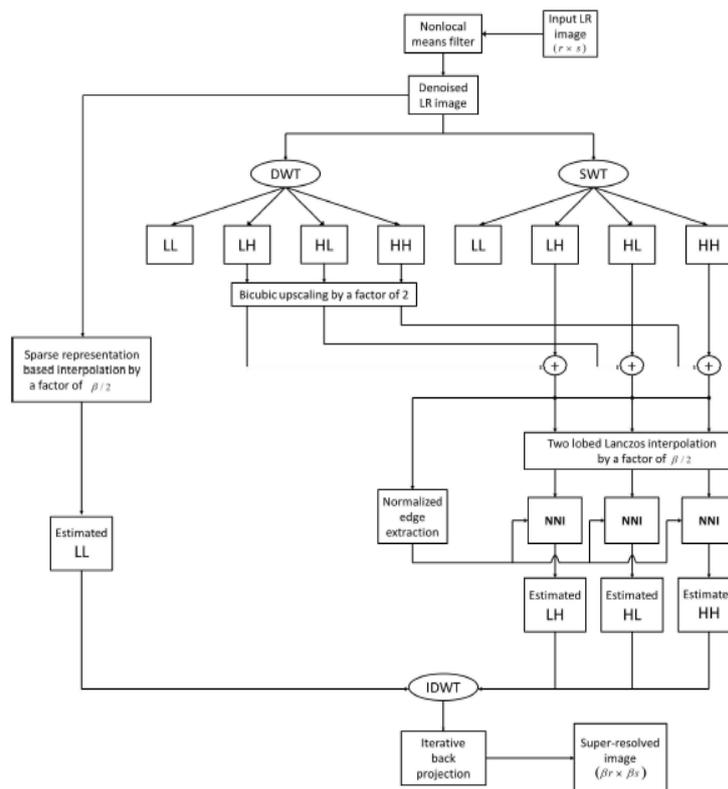


Figure 2. Block diagram of the proposed wavelet based SR technique

In general, the LL subband is considered as the low frequency version of the denoised LR image, since it is obtained by low pass filtering the input image. So, the LL subband does not contain any valid edge information. Hence we have processed only LH, HL and HH subbands, leaving behind the LL subband. The authors in [10] replaced LL subband with the input LR image. As mentioned earlier, this process leads to non-uniform illumination in the reconstructed SR image. So, we first apply sparse representation based interpolation method by a factor of 2 on the denoised LR input image and then the result is used in place of the LL subband. Finally, the sparse recovered image and the three estimated high frequency subbands are combined using IDWT. The obtained SR image can be further improved by applying the back projection algorithm, which iteratively works on minimizing the SR reconstruction error. The block diagram of the proposed technique is outlined in Figure 2.

4. Results and Discussion

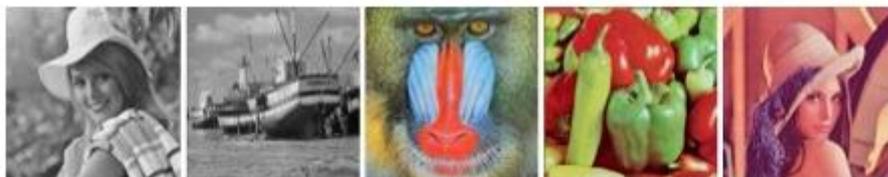


Figure 3. The LR test images used in our simulations

In this section, we present several experiments conducted to examine the performance of the proposed method. Figure 3 shows the LR test images used in our simulations. These images are produced by direct down sampling the original HR images by a factor of 4. Gray-level images (Elaine and Boat) as well as color images (Mandrill, Peppers and Lena) are

involved for testing. For color images we first transform the RGB color model into YCbCr color model and apply the SR reconstruction process only to the Y (luminance) channel, while the Cb and Cr channels are enlarged using plain Lanczos interpolation. Since human eye is highly sensitive to the changes in luminance than in color.

The conventional techniques: bicubic interpolation (Bicubic); Lanczos interpolation (Lanczos); edge guided directional filtering (EGDF) [3] and the state-of-the-art techniques: sparse representation model (SRM) [6]; sparse mixing estimators (SME) [5]; DWT based image enhancement (DWT) [13]; image SR using DWT and SWT (DWT-SWT) [10]; wavelet domain interpolation via sparse estimators (DWT-Sparse) [11] are compared with the proposed SR method. The above SR approaches are implemented using Matlab software which run on a system with an Intel 2.3GHz CPU and a 4GB RAM.

Table 1. PSNR (dB) and SSIM indices of the reconstructed images

Method	Image				
	Elaine	Boat	Mandril	Peppers	Lena
Bicubic	29.07	26.79	27.93	28.21	28.82
	0.923	0.841	0.758	0.883	0.943
Lanczos	29.24	26.41	28.01	28.62	28.85
	0.922	0.839	0.755	0.885	0.877
EGDF [3]	18.84	19.22	20.72	20.66	21.26
	0.959	0.916	0.830	0.899	0.931
SRM [6]	28.77	26.32	27.55	27.79	28.26
	0.875	0.931	0.835	0.902	0.872
SME [5]	33.12	28.88	30.35	32.10	32.89
	0.945	0.842	0.858	0.905	0.946
DWT [13]	29.12	29.48	29.99	30.21	33.01
	0.922	0.841	0.861	0.899	0.898
DWT-SWT [10]	34.13	29.92	30.11	32.39	33.12
	0.965	0.889	0.848	0.910	0.949
DWT-Sparse [11]	32.44	30.52	30.25	31.75	33.25
	0.973	0.941	0.851	0.905	0.941
Proposed method	34.57	30.77	30.33	32.45	33.79
	0.966	0.949	0.867	0.907	0.955

To denoise the input LR image using nonlocal means filter, we choose 5×5 pixels neighborhood Q and the range weight, $\sigma_R = 2$. The subband decomposition is performed using Biorthogonal (Bior 1.1) wavelet function. The LR image undergoes sparse representation based interpolation by a factor of 2. We choose $\lambda = 0.1$ for computing the sparse vector α^* in (1).



Figure 4. Reconstructed results of Elaine, Boat and Lena images using different methods: (a) Bicubic interpolation; (b) DWT [13]; (c) DWT-Sparse [11]; (d) Proposed method

For qualitative assessment, we have used the objective metrics: peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM). These metric values are tabulated in Table 1. In addition, the qualitative assessment is carried out in Figure 4 using visual quality as a subjective metric. By looking at Table 1 and Figure 4, we can notice the outstanding performance of the proposed SR technique.

5. Conclusion

In this paper, we have presented an effective image SR algorithm using the sparse signal recovery and wavelet domain interpolation. The nonlocal means filter in the preliminary stage makes our algorithm more robust to noise. The information loss due to the translational invariance and inferior directionality of the DWT is corrected using the SWT. In parallel, the sparse recovered image effectively serves as the estimated LL subband for the IDWT operation. Compared to the existing SR approaches, the proposed technique has the advantage of preserving the edge information. Our simulation results demonstrated that the proposed method generated qualitatively better images with high PSNR and SSIM values.

References

- [1] W Liang, S Chao-xuan, H Qiang, H Zhuang-zhi, W Yong-lei. Super-resolution imaging realization of costas signal. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2014; 12(2): 1374-1384.
- [2] X Li, MT Orchard. New edge-directed interpolation. *IEEE Transactions on Image Processing*. 2001; 10(10): 1521-1527.
- [3] D Zhang, X Wu. An edge-guided image interpolation algorithm via directional filtering and data fusion. *IEEE Transactions on Image Processing*. 2006; 15(8): 2226-2238.
- [4] X Zhang, X Wu. Image interpolation by adaptive 2-d autoregressive modeling and softdecision estimation. *IEEE Transactions on Image Processing*. 2008; 17(6): 887-896.
- [5] S Mallat, G Yu. Super-resolution with sparse mixing estimators. *IEEE Transactions on Image Processing*. 2010; 19(11): 2889-2900.
- [6] J Yang, J Wright, TS Huang, Y Ma. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*. 2010; 19(11): 2861-2873.
- [7] R Zeyde, Elad, M Protter. *On single image scale-up using sparse-representations*. In Curves and Surfaces. Springer. 2012: 711-730.
- [8] A Ranga, G Suryanarayana. A novel approach for single image super resolution by sparse signal representation. *Indian Journal of Science and Technology*. 2015; 8(S2): 74-77.
- [9] W Dong, L Zhang, G Shi, X Li. Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*. 2013; 22(4): 1620-1630.
- [10] H Demirel, G Anbarjafari. Image resolution enhancement by using discrete and stationary wavelet decomposition. *IEEE Transactions on Image Processing*. 2011; 20(5): 1458-1460.
- [11] H Chavez-Roman, V Ponomaryov. Super resolution image generation using wavelet domain interpolation with edge extraction via a sparse representation. *Geoscience and Remote Sensing Letters, IEEE*. 2014; 11(10): 1777-1781.
- [12] M Agrawal, R Dash. *Image resolution enhancement using lifting wavelet and stationary wavelet transform*. In Electronic Systems, Signal Processing and Computing Technologies (ICESC), 2014 International Conference on IEEE. 2014: 322-325.
- [13] G Anbarjafari, H Demirel. Image super resolution based on interpolation of wavelet domain high frequency subbands and the spatial domain input image. *ETRI journal*. 2010; 32(3): 390-394.
- [14] K Nasrollahi, TB Moeslund. Super-resolution: a comprehensive survey. *Machine vision and applications*. 2014; 25(6): 1423-1468.
- [15] Y Zou, X Liang, T Wang. Visible and infrared image fusion using the lifting wavelet. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(11): 6290-6295.