Single Image Super Resolution Using Sparse Image and GLCM Statistics as Priors

Subrahmanyam Ravishankar,Nagadastagiri Reddy.C, *Member, IAENG,* and M.V.Joshi

***Abstract*—Sparse representation of images finds many applications in image processing and computer vision. Recently various attempts have been made to regularize the ill-posed inverse problem of motion free image super resolution using sparse representation of low resolution image patches. However the proposed method in this paper is different from the previous approaches reported in the literature in terms of method of dictionary training and feature extraction from the trained data base images. Gray Level Co-Occurrence Matrix(GLCM) is a proven method for extracting image statistical features, which are used mainly for image classification, segmentation etc. In the present work we have extracted GLCM parameters for regularization of the data fitting term of the cost function of the image super resolution model. We used Matching Optimal Directions (MOD) algorithm[1] for obtaining high resolution and low resolution dictionaries from training image patches and seek the sparse representation of low resolution input image patch using low resolution dictionary and then obtain high resolution image patch from high resolution dictionary. The results of the proposed algorithm showed visual, PSNR, RMSE, and SSIM improvements over other super resolution methods.**

***Index Terms*—GLCM,DCT, OMP, MOD.**

I. Introduction

IMAGE super resolution (SR) is an active area of research as physical constraints limit image quality in many imaging applications. These imaging systems yield aliased and under sampled images if their detector array is not sufficiently dense. This is particularly true for infrared imagers and some charge-coupled device cameras. Super resolution is also crucial in satellite imaging[3], medical imaging, and video surveillance. There are several approaches for image super resolution such as bi-cubic interpolation, B-spline interpolation, etc. These approaches produce overly smooth images which lack high frequency components and thus blur the edge information of the reconstructed image. Other conventional approaches like multi frame super resolution create high resolution images employing more than one low resolution image. These methods use multiple low resolution images of the same scene, which are aligned with sub pixel accuracy to generate high resolution image. Recently, single frame image super resolution algorithms have been successfully employed for image super resolution. These methods use only one low resolution (LR) sample image to super resolve the LR image. Freeman et. al. proposed an example based super resolution technique[2]. They estimate missing high frequency details by interpolating the input low resolution image into a desired scale. The super resolution is performed by the nearest neighbor based estimation of high frequency patches corresponding to the patches of the low frequency input image. Super resolution of images is an ill-posed inverse problem due to unknown blurring operators and insufficient number of low resolution images. Different regularization methods have been proposed for providing solutions to this problem[4], [5], [6]. Yi Ma et. al. [9] proposed a super resolution algorithm based on sparse representations. The authors sparsely represents given low resolution image patches using low resolution dictionary and use these coefficients to construct high resolution patches from high resolution dictionary. It is also reported that the authors trained both high resolution and low resolution dictionaries using dual tree complex wavelet transform (DT-CWT) from a set of training images. Elad et. al. [10] have proposed another dictionary based technique for super resolution using sparse representation. The authors used k-means single value decomposition (K-SVD) algorithm for dictionary training. In our proposed approach Matching Optimal Directions (MOD) algorithm have been used for dictionary training, and Orthogonal Matching Pursuit (OMP) algorithm for sparse representations. We find that MOD algorithm is less complex than K-SVD algorithm and further by using regularization methods our results proved to be superior to the previous methods. From a set of low resolution (LR) images and corresponding high resolution images (HR) we extracted patches. MOD algorithm is applied on these patches to create low resolution dictionary (A_l) and high resolution dictionary (A_h). Low resolution (LR) test image patches are sparsely represented using A_l and using these sparse coefficients we obtained high resolution patches from A_h. We concatenate these patches to get high resolution image. Further we extract GLCM parameters of this image, for regularization. Gradient descent optimization algorithm has been used to minimize the cost function.

II. Image Formation Model

Let the original high resolution image be represented as $y_h \in R^N$, blur operator as $B : R^N \to R^N$ and down sampling operator as $D : R^N \to R^M$ where $M < N$, and $z_l \in R^M$ is defined as low resolution version of the original high resolution image. The observed image can be modeled as

$$z_l = BDy_h + n \tag{1}$$

Where **n** is an i.i.d Gaussian noise of mean zero and variance σ_n^2. The problem is to find high resolution estimate $\hat{y_h}$ for a given low resolution image z_l such that $\hat{y_h}$ is similar to y_h. We can obtain $\hat{y_h}$ by minimizing $\|BDy_h - z_l\|^2$, since BD is rectangular matrix we cannot find exact solution to

Manuscript received march 23, 2011; revised april 15, 2011.

Subrahmanyam Ravishankar is currently professor in the department of ECE, Amrita ViswaVidya Peetham University, Bangalore campus, India. e-mail:s_ravishankar@blr.amrita.edu.

Nagadastagiri Reddy.C is a student in the department of ECE, Amrita ViswaVidya Peetham University, Bangalore campus, India e-mail: dastagiri.c@gmail.com.

M.V.Joshi is currently professor in the department of ICT, DAIICT, Gandhinagar, India. e-mail:mv_joshi@daiict.ac.in.

above problem and there are infinitely many solutions for equation (1). Existing approaches for image super resolution use various priors such as textures, edges, and curves etc for providing solution to then above equation(1) as discussed in [7], [8].

III. Sparse Representation

Sparse representation accounts for most or all information of a signal with a linear combination of a small number of elementary signals called atoms in an over complete dictionary. Sparse representation is widely used in the areas of compressed sensing, image super resolution and face hallucination. The objective of sparse representation of a data vector $y \in R^m$ (for example a patch of an image) using a small number of non zero components in a sparse vector $x \in R^n$ under the linear generative model.

$$z_l = Ax + n \tag{2}$$

Where the full low rank dictionary $A \in R^{m\times n}$ may be over complete $(n > m)$, and the additive noise is assumed to be Gaussian with zero mean and variance σ_n^2.

IV. Proposed Algorithm

The proposed algorithm has two parts. Part-1 of the algorithm (Training phase) constructs dictionaries of low resolution and high resolution data sets, and Part-2 of the algorithm (reconstruction phase) super resolves the given input low resolution test image using the results of part-1.

A. Training Phase

The following steps are performed .

i. Training set construction: In this phase an image data bank is formed that consists of several high resolution images $\{y_h^j\}_j$ and the corresponding low resolution images $\{z_l^j\}_j$ which are blurred, down sampled versions of the high resolution images. $\{z_l^j\}_j$ are further interpolated to the size of high resolution image y_h.

ii. Feature extraction: After constructing the training set, low frequency components of the high resolution image that correspond to low frequency non- aliased components of the low resolution image are removed by computing Discrete Cosine Transform (DCT) and retaining coefficients which corresponds to high frequency information. The reason for this feature extraction is to focus the training on characterizing the relation between the low resolution patches and the edges, texture content within the corresponding high resolution ones[10]. Patches of size nxn are formed from low and high resolution images resulting in the data set $\{p_h^k, p_l^k\}_k$.

iii. Dictionary Training: MOD algorithm is used for training the dictionary. In this step MOD algorithm is employed on low-resolution image patches $\{p_l^k\}_k$ for training low resolution dictionary using the equation (3).

$$A_l, q^{(k)} = {}^{argmin}_{A_l,q^k} \sum_k \|p_l^k - A_l q^k\|^2 s.t. \|q^k\| \le L \tag{3}$$

High resolution dictionary A_h, is formed to recover high resolution patches by approximating them as $\{p_h^k = A_h q^k\}$. The dictionary A_h is constructed such that this approximation is as exact as possible, thus the dictionary is defined as the one which minimizes the mean approximation error, i.e.,

$$A_h = {}^{argmin}_{A_h} \sum_k \|p_h^k - A_h q^k\|_2^2 = {}^{argmin}_{A_h} \|P_h - A_h Q\|_F^2, \tag{4}$$

Where the matrix P_h is constructed with high resolution patches $\{p_h^k\}_k$ as its columns, and Q contains $\{q^k\}$ as its columns. The solution to the above problem is given by following pseudo-inverse expression[10] (given that Q has full low rank).

$$A_h = P_h Q^+ = P_h Q^T (QQ^T)^{-1} \tag{5}$$

The construction of high resolution and low resolution dictionaries A_h, A_l concludes training phase of super resolution algorithm.

B. Reconstruction Phase

In this phase our task is to super resolve the given low resolution image z_l using the following steps.

i. Obtaining high resolution estimate of z_l: The low resolution image is scaled by a factor of D using hermit interpolation resulting in $y_l \in R_N$. High frequency components are extracted from y_l using DCT, then patches $\{p_l^k\}_k$ are obtained from the extracted image. OMP[1] algorithm is applied to the patches $\{p_l^k\}_k$ and dictionary A_l, allocating L atoms to represent each patch for finding sparse representation vectors $\{q^k\}_k$. The sparse representation vectors $\{q^k\}_k$ are multiplied with high resolution dictionary A_h and the approximated high resolution patches $\{A_h q^k\}_k = \{p_h^k\}_k$ are obtained. The high resolution estimate $\hat{y_h}$ is obtained from $\{p_h^k\}_k$ by solving the following minimization problem with respect to $\hat{y_h}$

$$\hat{y_h} = {}^{argmin}_{A_h} \sum_k \|R_k(\hat{y_h} - y_l) - p_h^k\|_2^2 \tag{6}$$

R_k is patch extraction operator. Above problem states that patches from the image $\hat{y_h} - y_l$ should be as close as possible to the approximated patches, this problem has a least squares solution given by

$$\hat{y_h} = y_l + [\sum_k R_k^T R_k]^{-1} \sum_k R_k^T p_h^k \tag{7}$$

The term $R_k^T p_h^k$ positions the high resolution patch in the k^{th} location and term $R_k^T R_k$ is diagonal matrix that weighs each pixel in the high resolution outcome. Above equation can be simply explained as putting p_h^k in proper locations and adding y_l to them for getting high resolution estimate.

ii. Regularizing high resolution estimate: The high resolution estimate is regularized by solving thenfollowing cost function

$$C = \sum_{i,j} \|z_{l(i,j)} - f_{(i,j)}\|^2 - \|F_{hom(i,j)} - f_{hom(i,j)}\|^2 \tag{8}$$

Where z_l is the low resolution image, $f_{hom(i,j)}$ is homogeneity of interpolated image and $F_{hom(i,j)}$ is homogeneity of high resolution estimate. The above cost function is minimized using gradient descent optimization. GLCM is defined as the joint probability of occurrence of two gray

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

level values at a given offsets (in terms of distance and direction). For a given nxm imageI GLCM is computed as

$$G_{d,\theta} = \sum_{x=1}^{n}\sum_{y=1}^{m} 1, for I_{x,y} = i \& I_{x',y'} = j, \; and \, G_{d,\theta} = 0, else \tag{9}$$

where x' and y' are calculated from x, y and θ.$1 \leq x' \leq n, 1 \leq y' \leq m$. Homogeneity and correlation are calculated from GLCM, Homogeneity is the measure of closeness of the distribution of elements in the GLCM to the GLCM diagonal and Correlation measures the joint probability occurrence of the specified pixel pairs. In equation (8) correlation can also be used in place of homogeneity.

V. Experimental Results

We constructed a training set with 30 high resolution images and corresponding low resolution images by down sampling. High frequency components are extracted from the training set using DCT, and then we extracted 7860 high resolution and low resolution non-overlapping patches from training set. MOD algorithm is applied on the low resolution patches for constructing A_l, pseudo-inverse expression is used to calculate A_h. We applied our algorithm on several test images of similar classes as used by Elad .et.al [10] and Yi Ma .et.a [9]. We computed various metrics such as PSNR (peak signal to noise ratio), RMSE (root mean squared error) and SSIM (structural similarities) for our super resolved images. These results were compared with standard interpolation methods and the results obtained by Elad .et.al [10] and Yi Ma .et.al [9]. It is found that our approach gives better results than the above methods. The results are tabulated in table1,table2 and table3. Fig1 gives visual comparision of the results of the proposed methods with other listed methods.

VI. Conclusion

We compared our results with the results obtained by Yi Ma et. al. [9], Elad et. al.[10]. Our simulated results indicates better PSNR perfomance. We modified the method used by Elad et.al. [10] by employing MOD algorithm for dictionary training instead of KSVD algorithm, and DCT was used for feature extraction. DCT is simpler and effective because it eliminates patch concatenation and dimensionality reduction. We used very less number of training patches as compared to Elad et. al. [10]. and got better results. Still the efficiency of algorithm can be improved by optimizing thresholds for DCT and extracting patches with overlapping.

VII. Acknowledgement

The authors wish to thank Ron Rubinstein, research scholar, Technion Israel Institute of Technology for generously sharing matlab code for MOD and OMP algorithms.

References

[1] Michael Elad ., "Sparse and redundant representations: from theory to applications in signal and image processing,," *Springer publications* 1^{st} *edition.* ,2010.

[2] W. Freeman, T. Jones, E. Pasztor., "Example based super resolution," Computer graphics and Applications. , vol.22, no.2,pp.56-65,2002

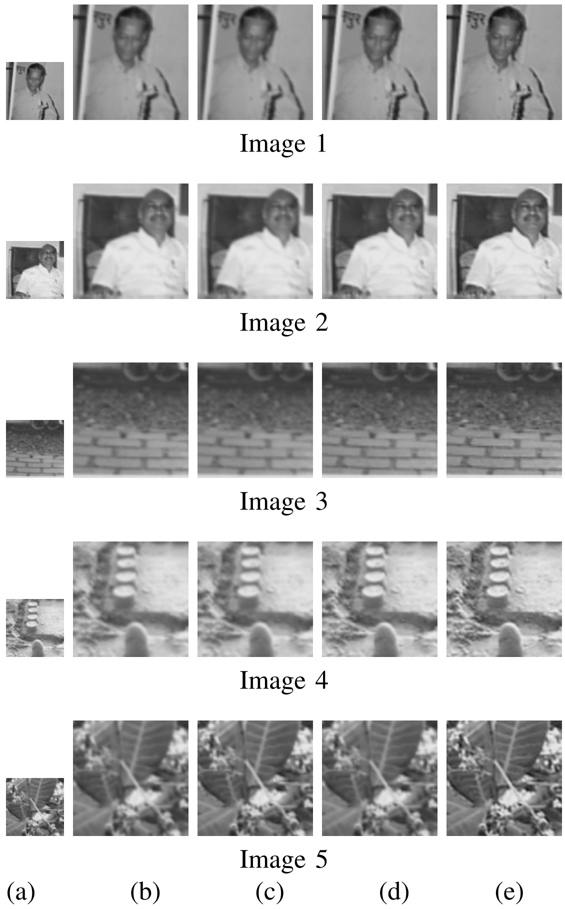

Fig. 1. (a) LR Image, (b) Mitchell Interpolation, (c) B-spline Interpolation, (d) Bell Interpolation, (e) Proposed Method.

TABLE I
RMSE comparison.

Image No.	RMSE			
	Mitchell interpolation	B-spline interpolation	Bell interpolation	proposed algorithm
1	12.2838	12.7950	11.6912	11.311
2	12.1617	12.6742	11.5603	11.1650
3	9.9298	10.2821	9.6024	9.4114
4	14.7493	15.2244	14.2904	14.1329
5	17.2603	18.0352	16.3310	15.6001

TABLE II
PSNR comparison.

Image No.	PSNR			
	Mitchell interpolation	Yi Ma et. al.	Elad et. al.	Proposed algorithm
1	33.20	33.10	33.50	38.89
2	33.10	33.10	33.50	33.72
3	24.50	24.80	25.00	28.66
4	24.76	24.80	25.00	30.51
5	28.00	28.20	28.40	31.27

TABLE III
SSIM comparison.

Image No.	SSIM			
	Mitchell interpolation	B-spline interpolation	Bell interpolation	Proposed algorithm
1	0.8750	0.8613	0.8911	0.9032
2	0.8435	0.8267	0.8608	0.8787
3	0.7808	0.7569	0.8092	0.8251
4	0.7395	0.7171	0.7663	0.8064
5	0.7547	0.7250	0.7896	0.8275

[3] H.Y.Liu, Y.S.Zhang., "Study on the methods of super resolution image reconstruction," *The International archives of photogrammetry, remote sensing and spatial information sciences.* , Vol.XXXVII .part B2. Beijing,2008

[4] R.C. Hardie, K.J. Barnard, and E.A. Armstrong., "Joint MAP registration and high resolution image estimation using a sequence under sampled images," TIP. ,1997

[5] S. Farisu, M.D. Robinson, M. Elad, and P. Milanfar., "Fast and robust multiframe super resolution," TIP. ,2004

[6] M.E. Tipping and C.M. Bishop., "Bayesian image super resolution," . Neural Information Processing Systems. ,2003

[7] L.C. Pickup, S.J. Roberts, and A. Zisserman., "A sampled texture prior for image super resolution," . Nueral Information Processing Systems. ,2003

[8] K. Rudraraju and M.V. Joshi., "Nonhomogeneous AR model based prior for multiresolution fusion," . IEEE international Geosciences and Remote Sensing Symposium. ,2009

[9] Jianchao Yang, John Wright, Thomas Huang, Yi Ma., "Image super resolution via sparse representation," . IEEE TIP. , 2010

[10] Roman Zeyde, Michael Elad and Matan Protter., "On Single Image Scale-Up using sparse representations," *Curves and Surfaces, Avignon-France.* ,2010

ISBN: 978-988-19251-4-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)