ISSN: 2321-8169
2497 – 2502

Single image super resolution using compressive K-SVD and fusion of sparse approximation algorithms

Aneesh G Nath
Department of Computer Science and Engineering
TKM College of Engineering
Kollam, India
aneeshgnath@yahoo.co.in

Retheesh V V
Department of Computer Science & Engineering
NITK Surathkal, Manglore
Karataka, India
rethboby@gmail.com

Abstract— Super Resolution based on Compressed Sensing (CS) considers low resolution (LR) image patch as the compressive measurement of its corresponding high resolution (HR) patch. In this paper we propose a single image super resolution scheme with compressive K-SVD algorithm(CKSVD) for dictionary learning incorporating fusion of sparse approximation algorithms to produce better results. The CKSVD algorithm is able to learn a dictionary on a set of training signals using only compressive sensing measurements of them. In the fusion based scheme used for sparse approximation, several CS reconstruction algorithms participate and they are executed in parallel, independently. The final estimate of the underlying sparse signal is derived by fusing the estimates obtained from the participating algorithms. The experimental results show that the proposed scheme demands fewer CS measurements for creating better quality super resolved images in terms of both PSNR and visual perception.

Keywords*- super resolution; compressive K-SVD;OMP;sparsity.*

*****

I. INTRODUCTION

Super-resolution (SR) is the process of generating single high resolution (HR) image from one or more low resolution (LR) images. In traditional methods of SR the low resolution images that were captured with sub-pixel accuracy are used to solve the missing high-frequency information. But it is more challenging to recover this information from a single, low-resolution image. In many applications only a single low-resolution image is available and the Single image super-resolution (SISR) problem is particularly important in those situations. Compressed sensing (CS) is an emerging data acquisition technique which overcomes the limitations of Shannons sampling theorem. The motivating fact behind CS is that many natural signals are sparse or approximately sparse in a certain basis like wavelet, fourier etc. In many emerging applications, the abundance of data generated by the sensing systems due to high sampling rate demands data compression before storage or transmission.

Compressed Sensing combines the sampling and the compression into a single process. CS data acquisition technique enables the reduction in the number of measurements required for recovery of sparse signals or compressible signals which is sparse on some suitable basis. Reconstruction of the signals from CS measurement is done using greedy or relaxation based algorithms. In a CS based image acquisition system it acquires less number of random linear measurements (pixels at a subset of sampling lattice) without first collecting all the pixel values. Proper preprocessing technique will enable the reconstruction of image from this incomplete data. The resulting image is an LR image which is suitable for reconstruction of its original HR image. In compressive sensing based single image super resolution, a low resolution input image plays the role of the compressive measurement of its corresponding high resolution image, and a proper dictionary which represents the high resolution image sufficiently sparse will make an accurate recovery of high resolution image. The problem of SISR is to obtain HR image X from its degraded Low-Resolution (LR) version Y, represented as

$$Y = HBX + v \qquad (1)$$

where H, B and v represent the downsampling operator, blurring operator and the additive noise respectively.

In this recovery problem, we use CS based reconstruction method based on dictionary learning to generate the HR image. CKSVD an algorithm for learning a dictionary based on a given set of CS measurements, which is a generalization of the well-known K-SVD algorithm. More precisely, this algorithm is an iterative approach that alternates between sparse coding and dictionary update steps to minimize the error in representation of the CS measurements. The sparse coding stage is performed by fusing sparse approximation algorithms like OMP, SP etc.

This paper is organized as follows. Section 2 describes the related works reported in the literature, Section 3 describes about the proposed super-resolution scheme including the dictionary training and reconstruction phases with the preprocessing step. Experimental results are shown in Section 4 and conclusions are drawn in Section 5.

II. LITERATURE SURVEY

SR reconstruction has been one of the most active research areas since the 1984 by seminal work of Tsai et al.[1] Many techniques have been proposed over the last three decades from frequency domain approach of Borman et al.[2] to spatial domain approach of Sung et al.[3], and from signal processing perspective to machine learning perspective. Conventional approach in super resolution is to generate a SR image from multiple low-resolution input images which are registered and

ISSN: 2321-8169
2497 – 2502

reconstructed into a high-resolution image using constraints (e.g. Bilateral Total Variation[13] and Huber Markov Random Fields (MRF)[12]), with the help of maximum a-posteriori (MAP) like regularization technique.

Single Image SR (SISR) has mainly two kinds of methods.

1) Reconstruction-based super-resolution[11],[14],[16] without training process but with the help of well defined constraints for the target high-resolution image.

2) Learning-based super-resolution algorithms[18]-[21] make use a dictionary which is trained and tested using training set of images.

Interpolation based algorithms reconstruct the image details by interpolating the LR input image and enhance the edges by making it sharper. The standard methods in this category are bilinear and bicubic interpolation. Another well-known algorithm is Back Projection[17]. This method gradually sharpens the edges while it iterates through the image. The Dai et al.[16]'s approach is also an example of above category. In order to enforce their continuity it extracts the edges of the image and blends them with the interpolated result. Fattal[14] uses edge statistics to reconstruct the missing high frequency information. The example-based super-resolution method[19] try to recover the lost image details with the help of a database consisting of co-occurrence examples from a training set of HR and LR image patches. But the main drawback of example-based algorithms is that they heavily rely on the similarity between the training set and the test set. The sparse representation method based on CS[4] theory exploits the linear relationships among high-dimension signals and uses it to recover HR information from their low-dimension projections.

Yang et al.[5],[6] proposed that the idea of sparsity can be used instead of working directly with the patch pairs sampled from high resolution and low resolution images. The method [5] trained overcomplete dictionaries by low-resolution and high resolution images for representing them as sparse vectors to capture the co-occurrence prior to improve the speed and the robustness. [6] leant a compact representation for these patch pairs and this method achieved significant performance improvement. Later Zeyde et al.[7] modified approach proposed by [6], which is much faster and efficient than [6]. Zhang et al.[15] proposed a novel image super-resolution method via dual dictionary learning and sparse representation, which consists of the main dictionary learning and the residual dictionary learning to recover main high-frequency(MHF) and residual high-frequency(RHF) respectively.

III. PROPOSED RECOSTRUCTION SCHEME

Like other learning based super resolution methods, our approach also consists of dictionary learning phase and image recostruction phase. We adopt the method proposed by [7] which is proved to be faster and efficient. The following section describes about CKSVD[8] algorithm with fusion of OMP and SP[9] for sparse coding stage. The next section explains the overall super resolution scheme.

A. Compressive KSVD

Given a set of CS measurements for n training signals $\{m_i\}_{i=1}^{n}$ in $\mathbb{R}^m$, the goal is to find a dictionary $D \in \mathbb{R}^{p\times d}$ that leads tothe best possible representation for the original training signalsY under the strict sparsity constraint.

The universal CS measurement matrix is the random Gaussianmatrix with each entry drawn i.i.d from $\mathcal{N}(0,1)$. Let $E_i \in \mathbb{R}^{p\times m}$ denote the measurement matrix used for the i^{th}training signal. Then, each measurement vector can be written as follows

$$m_i = E_i^T y_i \in \mathbb{R}^m \text{ for } i = 1, 2, ..., n \qquad (2)$$

It is very important to consider different measurement matrices for different training signals here. Otherwise, by projecting all the training signals onto one low-dimensional random subspace, the original signal space is lost, and we will not be able to find the sparse representation model for the original training signals.

We assume that the dictionary model is D = BA where B is a fixed matrix and the matrix A is the atom representation dictionary. The matrix B can contain some prior knowledge about the training signals, e.g. principal components of them learned through another method, or it can be the identity matrix. In this model, the i^{th} column of A denoted by $a^{(i)}$ corresponds to the ith dictionary atom $d^{(i)} = Ba^{(i)}$, for $i = 1, 2, ..., d$. Therefore, each measurement vector m_i is

$$m_i = E_i^T y_i = E_i^T BAx^{(i)} \text{ for } i = 1, 2, ..., n. \qquad (3)$$

A generalization of the objective function in (1) for the case that we have access only to the CS measurements is to seek the dictionary D that leads to the best representation for the CS measurements, rather than the original signals, under the strict sparsity constraint. Our proposed objective function thus becomes a well-defined penalty term that measures the quality of our sparse representation in terms of the only information available from the original data. This leads to the criterion:

$$\min_{A,X} \sum_{i=1}^{n} \left\| m_i - E_i^T BAx^{(i)} \right\|_2^2 \text{ subject to } \forall_i, \left\| x^i \right\|_0 \le T \quad (4)$$

Clearly, for $E_i = I_{p\times p}$, $i = 1,\ldots, n$, this objective function reduces to the usual K-SVD[10] criterion. Solving the optimization problem will give us the atom representation dictionary A and the corresponding dictionary D = BA such that $y_i = Dx^i$ or $\left\| y_i - Dx^i \right\|_2 \le \varepsilon$ for some small ε.

In our proposed CK-SVD algorithm, the penalty term in (4) is minimized in a simple iterative approach, similar to K-SVD's,that alternates between sparse coding and dictionary update steps.

(i) Sparse coding step

In the sparse coding step, the penalty term in (4) is minimized with respect to a fixed A to find the best coefficient matrix X under the strict sparsity constraint. This can be written as

 ISSN: 2321-8169
2497 – 2502

$$\min_{X}\sum_{i=1}^{n}\left\| m_i-\Psi_i x^{(i)}\right\|_2^2 \text{ subject to } \forall_i, \left\| x^i \right\|_0 \le T \quad (5)$$

where $\Psi_i = E_i^T BA$ is a fixed equivalent dictionary for representation of the i[th] measurement vector m_i. This optimization problem can be considered as n distinct optimization problems for each measurement vector. We use fusion of two CS reconstruction algorithms: Orthogonal Matching Pursuit (OMP) and Subspace Pursuit (SP). Let T denote the actual support-set, and $\hat{T}^{(\mathrm{OMP})}$ and $\hat{T}^{(\mathrm{SP})}$ denote the support-sets estimated by OMP and SP respectively. Let $\hat{T}_{true}^{(\mathrm{OMP})} = T \cap \hat{T}^{(\mathrm{OMP})}$ and $\hat{T}_{true}^{(\mathrm{SP})} = T \cap \hat{T}^{(\mathrm{SP})}$ represent the sets of true atoms estimated by OMP and SP, respectively. Then we have $|T| = \hat{T}^{(\mathrm{OMP})} = \hat{T}^{(\mathrm{SP})} = K$, $0 \le |\hat{T}_{true}^{(\mathrm{OMP})}| \le K$ and $0 \le |\hat{T}_{true}^{(\mathrm{SP})}| \le K$ .The fact is that union set always contains at least as many true atoms as in the estimated support-set obtained from the best performing algorithm.

(ii) Dictionary update step

Assume that the $a^{(j)}, j \ne k$ and coefficients $x_T^j, j \ne k$ are fixed.

The goal is to update the k[th] dictionary atom, or equivalently $a^{(k)}$, and its corresponding coefficients sequentially for k = 1,…, d.
The penalty term in (4) can be written as

$$\sum_{i=1}^{n}\left\| m_i - E_i^T BAx^{(i)} \right\|_2^2$$
$$= \sum_{i=1}^{n}\left\| m_i - E_i^T B \sum_{j=1}^{d} a^{(j)} x_j^{(i)} \right\|_2^2$$
$$= \sum_{i=1}^{n}\left\| (m_i - E_i^T B \sum_{j\ne k}^{d} a^{(j)} x_j^{(i)}) - E_i^T Ba^{(k)} x_k^{(i)} \right\|_2^2$$
$$\sum_{i\in T_k}\left\| M_k^{(i)} - E_i^T Ba^{(k)} x_k^{(i)} \right\|_2^2 + \sum_{i\notin T_k}\left\| M_k^{(i)} \right\|_2^2 \quad (6)$$

where $x_k^{(i)}$ is a scalar corresponding to the coefficient of the k[th] dictionary atom in the representation of m_i with respect to Ψ_i, I_k is a set of indices of measurement vectors using the k[th] dictionary atom defined as follows

$$I_k = \{ \mathrm{i} \mid 1 \le \mathrm{i} \le \mathrm{n}, x_k^{(i)} \ne 0 \} \quad (7)$$

and $M_k^{(i)}$ is the representation error for the i[th] measurement vector when the k[th] dictionary atom is removed. Because the E_is are distinct this problem cannot be solved by SVD as before. However, the penalty term in (6) is a quadratic function of $a^{(k)}$ and the minimize is obtained by setting the derivative of the penalty term with respect to $a^{(k)}$ equal to zero to obtain

$$B^T(\sum_{i\in T_k}(x_k^{(i)})^2 E_i E_i^T)Ba^{(k)} = B^T(\sum_{i\in T_k}(x_k^{(i)}) E_i M_k^{(i)}) \quad (8)$$

Defining the matrix G_k and the vector b_k as follows

$$G_k \triangleq B^T(\sum_{i\in T_k}(x_k^{(i)})^2 E_i E_i^T)B$$
$$b_k \triangleq B^T(\sum_{i\in T_k}(x_k^{(i)}) E_i M_k^{(i)}) \quad (9)$$

we thus find that the update of the k[th] column of matrix A, $a^{(k)}$ is obtained by

$$a^{(k)} = G_k^+ b_k \quad (10)$$

Given the new $a^{(k)}$ the optimal where $x_k^{(i)}$ for each $i \in I_k$ is given by least squares as follows. It is clear that the support of the coefficient matrix X is preserved as in the K-SVD algorithm.

$$x_k^{(i)} = \frac{\left\langle M_k^{(i)}, E_i^T Ba^{(k)} \right\rangle}{\left\| E_i^T Ba^{(k)} \right\|_2^2} \quad (11)$$

B. Training and Reconstruction scheme using dictionary

The overall super resolution scheme is divided into two stages named Dictionary learning phase using CKSVD and Image synthesis phase.

(i) Dictionary learning phase

This stage starts with collecting a set of High resolution images H_{ORG} from the training set. Since we assume our algorithm to work with samples from a Compressed Sensing based acquisition system we take downsampled and blurred version of these images as input.
To incorporate with the steps followed by [7], we consider the upscaled LR image denoted by L_{ORG} as the degraded version of the high resolution image H_{ORG}. Then high frequency features are extracted by subtracting L_{ORG} from H_{ORG}. As next step local patches are extracted from H_{ORG} and L_{ORG} considering only locations $k \in \Omega$ forming the data-set P = $\{p_l^k, p_h^k\}_k$. LR image L_{ORG} is filtered by using R high pass filters such as Laplacian in order to extract local features that correspond to their high-frequency content. The last step before turning to the dictionary learning stage is reducing the dimension of the input low-resolution patches, by Principal Component Analysis (PCA) algorithm. Next, the CKSVD dictionary training is applied to the set of patches.

Low Resolution Dictionary Training:
The dictionary training stage starts with the low-resolution patches $\{\hat{p}_l^k\}_k$ extracted from low resolution image. As the result of the applying CKSVD dictionary learning algorithm to these patches, the dictionary $L_D \in R^{n\times m}$ is created.

$$L_D, \{q^k\} = \arg \min_{L_D, \{q^k\}} \left\| p_l^k - L_D \{q^k\} \right\|^2 \ s.t \left\| q^k \right\|_0 \le L \forall k$$

. (12)

This training process also generates the sparse representation coefficients vectors q^k corresponding to the training patches $\{\hat{p}_l^k\}_k$ as a side product. Here $\| \ \|_0$ is the l_0 norm counting the nonzero entries of a vector.

High Resolution Dictionary Learning:

By approximating $p_h^k \approx A_h q^k$, for recovering p_l^k ,the already generated sparse representation vector q^k for the low resolution patch is multiplied by the high-resolution dictionary. The high resolution dictionary H_D can be found to get the correct approximation. So H_D is the dictionary matrix which minimizes the approximation error.

$$H_D = \arg \min_{H_D} \sum_k \left\| p_h^k - H_D \{q^k\} \right\|_2^2 \quad \text{i.e.}$$

$$H_D = \arg \min_{H_D} \left\| P_h - H_D Q \right\|_F^2 \tag{13}$$

The following Pseudo-Inverse expression is used for solving this problem.

$$H_D = P_h Q^+ = P_h Q^T (QQ^T)^{-1} \tag{14}$$

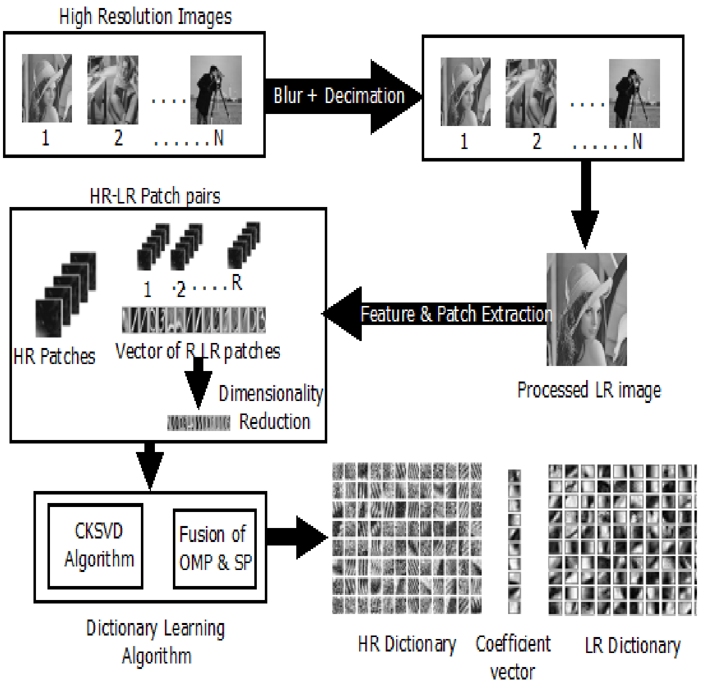

Fig. 1. Illustration of dictionary learning phase.

(ii) Image synthesis phase

In this phase the image is reconstructed using the trained dictionary to enhance the low-resolution image L_{ORG}. This image is assumed to be generated from a high-resolution image H_{ORG} by the same degradation operations (blur and scale-down) as in the training phase. For extracting the features this image is filtered using the same R high-pass filters that we have used in the training phase. Patches are extracted from these R images from locations $k \in \Omega$.

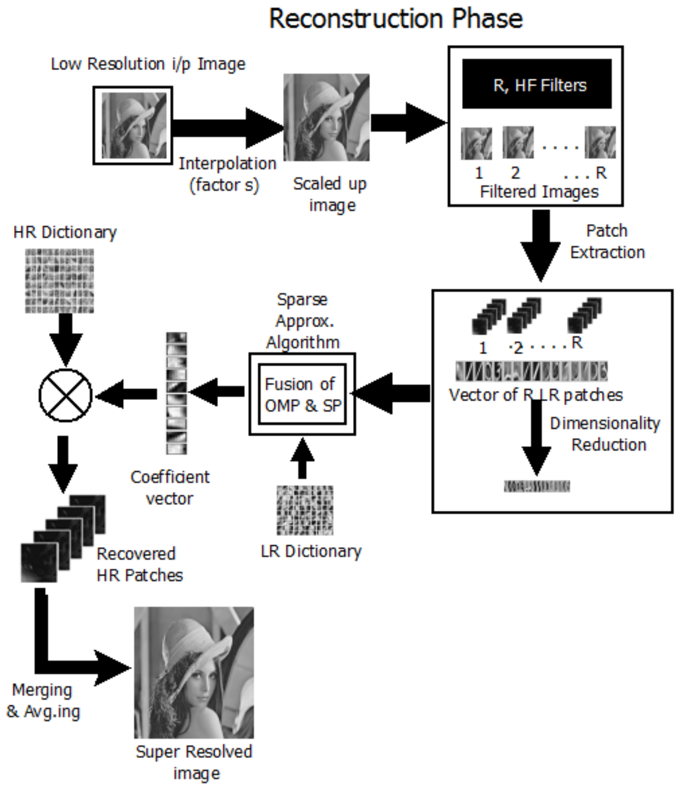

Fig. 2. Illustration of image synthesis phase

After the PCA based dimensionality reduction operation, the fusion of OMP and SP algorithms is applied to $\{\hat{p}_l^k\}_k$, allocating L atoms to their representation, and finding the sparse representation vectors $\{q^k\}_k$. The representation vectors $\{q^k\}_k$ are multiplied by the high-resolution dictionary H_D and the approximated high-resolution patches, $\{H_D q^k\}_k = \{\hat{p}_h^k\}_k$ are obtained. The final super-resolved image $\hat{H}_{ORG}$ is constructed from $\hat{p}_h^k$ by solving the following minimization problem with respect to $\hat{y}_h$:

$$\hat{y}_h = \arg \min_{\hat{y}_h} \sum_k \left\| R_k (\hat{y}_h - y_l) - \hat{p}_h^k \right\|_2^2 \tag{15}$$

Where R_k is the operator which extracts a patch of size $n \times n$ from the image in location k. ie. this problem states that extracted patches from the resulting difference image $\hat{y}_h - y_l$, should be as close as possible to the approximated patches

 ISSN: 2321-8169
2497 – 2502

$\left\{\hat{p}_h^k\right\}_k$. This problem has a closed-form Least-Squares solution,

$$\hat{H}_{ORG} = \hat{y}_h = y_l + \left[\sum_k \mathrm{R}_k^T \mathrm{R}_k\right]^{-1} \sum_k \mathrm{R}_k^T \hat{p}_h^k \qquad (16)$$

it is equivalent to putting $\left\{\hat{p}_h^k\right\}_k$ in their proper locations, averaging in overlap regions, and adding y_l to get the final image $\hat{H}_{ORG}$.

IV. EXPERIMENTAL RESULTS

The proposed super resolution scheme is implemented in MATLAB R2012a using CKSVD and fusion of OMP and SP algorithms, on Intel Core i5 – 2410M CPU at 2.30GHz with 4GB of RAM. In this framework the image is downscaled with scaling factor 2 before taking the in order to show the upscaling performance. The initial interpolation is done by Bicubic interpolation. Adding more and more images for training would lead to improved results. Extraction of features from the low resolution image is done with 4 filters that perform first and second horizontal and vertical derivatives: $f_1 = [1,-1] = f_2^T$ and $f_3 = [1,-2, 1] = f_4^T$. The patch size used is $n = 81$ (9×9), and the PCA results with a reduction from $4\times81 = 324$ dimensions to $n_l = 30$ dimensions. The dictionary training procedure applied 40 iterations of the CKSVD algorithm, with $m = 1000$ atoms in the dictionary, and allocating $L = 3$ atoms for each representation vector.

Fig.3. Visual comparison of the results.

Table 1. PSNR comparison

Image	Yang et al.	Zeyde et al.	Proposed method
Boat	25.2444	26.1259	26.4113
House	32.7651	35.3247	36.4401
Lena	28.3229	29.0562	30.2004
Peppers	26.9227	27.2600	27.3487
Man	25.8304	26.4995	27.6356
Barbara	21.8747	22.2540	23.0045
Fingerprint	22.0286	23.3575	24.4922
Couple	24.5410	25.1162	26.4198
Cameraman	21.4369	21.9221	22.2173

The results shown in figure 3 are the visual comparison of the proposed method with Yang et al. and Zeyde et al. methods which are using optimized implementation for KSVD and OMP algorithms . Table 1 shows the PSNR comparison of the same. By analysing the results it is obvious that incorporating CKSVD and fusion of OMP and SP improves the results in terms of PSNR and visual perceptions.

V. CONCLUSION

This paper presents an image super-resolution approach for creating high resolution images incorporating CKSVD and fusion of OMP and SP in dictionary learning. The results are give with quantitative and visual analysis. The results show that the proposed method is able to outperform the available dictionary learning schemes which are using KSVD and OMP like algorithms for creation of dictionary and sparse approximation.

REFERENCES

[1] R. Y. Tsai and T. S. Huang, "Multiple frame image restoration and registration", In *Advances in Computer Vision and Image Processing*, pages 317-339. Greenwich, CT: JAI Press Inc., 1984.
[2] Sean Borman and Robert L. Stevenson, "Super-resolution from image sequences - A review". In *Proceedings of the 1998 Midwest Symposiumon Circuits and Systems*, pages 374-378, 1998.
[3] Sung C. Park, Min K. Park, and Moon G. Kang, "Super-resolution image reconstruction: a technical overview", *IEEE Signal Processing Magazine*, 20(3):21{36., 2003.
[4] D. L. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[5] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches", in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

ISSN: 2321-8169
2497 – 2502

[6] J. Yang, J. Wright, T. Huang, and Y. Ma, “Image superresolution via sparse representation”, *IEEE Trans. on Image Processing*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[7] R. Zeyde, M. Elad, and M. Protter, “On Single Image Scale-Up using Sparse-Representations”, *Curves & Surfaces*, LNCS 6920, pp. 711–730, 2011.

[8] Farhad Pourkamali Anaraki, Shannon M. Hughes, “Compressive K-SVD”, Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, 26-31 May 2013 .

[9] Sooraj K. Ambat,Saikat Chatterjee,K.V.S. Hari,, “Fusion of Algorithms for Compressed Sensing”, IEEE Transactions on signal processing, Vol. 61, NO. 14, July 15, 2013.

[10] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation”, *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006

[11] Pradeep Sen, Soheil Darabi, “Compressive Image Super-resolution”, *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference*, pages 1235 – 1242, 1-4 Nov. 2009.

[12] R. C. Hardie, K. J. Barnard, and E. A. Armstrong, “Joint map registration and high-resolution image estimation using a sequence of undersampled images”, *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1621–1633, Dec. 1997.

[13] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, “Fast and robust multiframe super-resolution, *IEEE Trans. Image Process.”*, vol. 13, no.10, pp. 1327–1344, Oct. 2004.

[14] R. Fattal, “Image upsampling via imposed edge statistics”, *ACM Trans.Graph.*, vol. 26, no. 3, p. 95, 2007.

[15] Jian Zhang, Chen Zhao, Ruiqin Xiong, Siwei Ma, Debin Zhao, “Image super resolution via dual dictionary learning and sparse representation”, *IEEE Int. Symposium of Circuits and Systems (ISCAS)*, 2012.

[16] S. Dai, M. Han, W. Xu, Y. Wu, and Y. Gong, “Soft edge smoothness prior for alpha channel super resolution”, in *Proc. IEEE Conf. Comput.Vis. Pattern Class*, pp. 1–8, 2007.

[17] M. Irani and S. Peleg, “Motion analysis for image enhancement: Resolution, occlusion, and transparency”, *J. of Visual Communication and Image Representation*, vol. 4, no. 4, pp. 324–335, Dec 1993.

[18] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, “Learning lowlevel vision”, *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 25–47, 2000.

[19] Freeman, W.T., Jones, T.R., Pasztor, E.C., “Example-based super-resolution”. IEEE Comput. Graph. Appl. 22 (2), 56–65, 2002.

[20] J. Sun, N. N. Zheng, H. Tao, and H. Shum, “Image hallucination with primal sketch priors”, in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2, pp. 729–736, 2003.

[21] H. Chang, D.-Y. Yeung, and Y. Xiong, “Super-resolution through neighbour embedding”, in *Proc. IEEE Conf. Comput. Vis. Pattern Class.*, vol. 1, pp. 275–282, 2004.