

EDGE PRESERVING SINGLE IMAGE SUPER RESOLUTION IN SPARSE ENVIRONMENT

Srimanta Mandal and Anil Kumar Sao

School of Computing and Electrical Engineering
Indian Institute of Technology Mandi, India
srimanta_mandal@students.iitmandi.ac.in, anil@iitmandi.ac.in

ABSTRACT

Quality of an image is associated with edge of the image. It is important to preserve the edge of the image while deriving high resolution (HR) image from low resolution (LR) image, also known as super-resolution (SR) problem. This paper proposes an edge preserving constraint, which preserve the edge information of image by minimizing the differences between edges of LR image and the edges of the reconstructed image (down-sampled version), in sparse coding based SR problem. Partial edge evidences, derived using 1-D processing of image, are used separately in the constraints. The experimental results show that proposed approach preserves the edges of image as well as outperforms objectively the existing SR approaches.

Index Terms— Image Super-resolution, Sparse coding, Dictionary, Edge evidence.

1. INTRODUCTION

In many of image processing applications such as medical imaging, remote sensing and surveillance, high resolution (HR) image of region is crucial to characterize the region [1]. HR image may not be feasible often due to the limitations of imaging environment. This issue is addressed by deriving HR image from low resolution (LR) image, also known as super resolution (SR) problem [1] in image processing. Several approaches have been proposed in the literature to derive the HR image from LR image. These approaches use either several LR images of the same scene [1, 2] or single LR image of the scene [3, 4, 5, 6, 7]. The HR image, derived using the approaches based on the former strategy, are perceptually good provided the captured multiple images satisfy some specific conditions [1]. In many applications, we may not have multiple images of the scene and for this scenario, approaches based on latter strategy is very useful, where the missing information is adopted from several examples of HR image.

The SR problem can be formulated as

$$\mathbf{y} = \mathbf{S}\mathbf{B}\mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{y} \in \mathcal{R}^N$ is low resolution image that is produced by blurring operator \mathbf{B} and down-sampling operator \mathbf{S} in the original image $\mathbf{x} \in \mathcal{R}^M$, $N \ll M$. Here \mathbf{n} denotes the additive noise. Equation (1), can be seen as the problem of image restoration [8], is an ill-posed problem and some additive constraints are required to regularize the solution. Various constraints have been used in the literature [9, 10, 11], and each constraint has its own pros and cons. For example total variation (TV) model [10, 2] supports piecewise constant structure, so it tends to smooth out finer image details.

Recently, sparse coding/compressive sensing has become widely used approach in the area of signal processing because of its

capability to represent the signal at the rate well below Shannon/Nyquist's sampling criterion [12]. It has been used extensively in many image restoration algorithms [13, 14, 15, 16, 8]. In this approach, an image \mathbf{x} can be represented as linear combination of few columns of a matrix \mathbf{A} , defined as dictionary. Thus equation (1) can be written as

$$\mathbf{y} = \mathbf{S}\mathbf{B}\mathbf{A}\mathbf{c} + \mathbf{n}, \quad (2)$$

where \mathbf{c} is the sparse vector. Original HR image can be approximated if we have the sparse vector (i.e. $\mathbf{x} = \mathbf{A}\mathbf{c}$), and it is derived by solving the following criterion [17]:

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c}} \{ \|\mathbf{y} - \mathbf{S}\mathbf{B}\mathbf{A}\mathbf{c}\|_2^2 + \lambda \|\mathbf{c}\|_1 \}, \quad (3)$$

where λ denotes the Lagrange multiplier. Here, the sparsity of vector \mathbf{c} is approximated using l_1 -norm rather than l_0 -norm, because solving l_0 -norm is NP-hard problem and l_1 -norm minimization is the closest convex approximation of l_0 -norm minimization [18].

Due to the minimization of the term $\|\mathbf{y} - \mathbf{S}\mathbf{B}\mathbf{A}\mathbf{c}\|_2^2$ in equation (3), the finer details such as edges and textures of derived HR image are blurred. This issue is addressed by ASDS-AR-NL approach [8] using autoregressive and non-local mean constraint in SR problem. Results were further improved in [19] by adding another constraint that minimizes the differences between sparse vector of the degraded image and undegraded image, defined as sparse coding noise.

This paper proposes an additional constraint in equation (3), which preserve the edge information of image by minimizing the differences between edges of LR image and the edges of the reconstructed image (down-sampled version), in order to reduce the blurring of edges in approximated HR image. Edge is a vector quantity (it has direction as well as magnitude), and it is approximated by computing edge evidences along two orthogonal directions. We have used the partial edge evidences, computed by 1-D processing of image [20], separately in the constraints.

The organization of the paper is as follows: Section 2 describes the formation of dictionary which is crucial in sparse representation of images. The proposed edge preserving constraint in SR process is explained in Section 3. The experimental results are depicted in section 4, and section 5 summarizes the paper.

2. DICTIONARY FOR SR

Dictionary (\mathbf{A}), also known as sensing matrix, plays an important role in sparse representation of images. Several dictionaries have been proposed in the literature and can be categorized into two classes: (1) Analytic based dictionary (DCT, wavelet, etc.) and (2) learning based dictionary (MOD, K-SVD, PCA, etc.) [21]. Analytic dictionaries share the advantage of fast computation but they lack



Fig. 1. Edginess results of the *Bike* image - Left to right: Gray scale image, Vertical Edge Evidence, Horizontal Edge Evidence and Magnitude of Edge respectively.

adaptivity to image locality. This issue is addressed in the second class of dictionary by learning the details from several HR images, but they are not computationally as fast as compared to analytic based dictionary. However, they can represent image better than it's counterpart. Moreover it has been observed that sub-dictionaries created for patches can represent an image better than over-complete dictionary created for the whole image [8]. Hence, we have adopted the patch based dictionary learning as explained in [8], where some collected HR images are sliced into patches. Among those the smoothed patches are excluded and remaining are clustered using K-means clustering algorithm. Principal component analysis (PCA) has been applied to each cluster to learn patch based local dictionary. During SR process these dictionaries are selected adaptively.

3. EDGE PRESERVING SR

It is conjectured that the edge information of an image contains perceptually significant information of the image and also associated with quality of image. Hence, it is very important to incorporate the constraint of edge preservation during the generation of HR image. The work given in [16], uses the derivative features (edge features) to generate the SR image. But this method does not explain the significance of partial edge evidences while deriving HR image. In addition, the computation of derivative features are not efficient due to the discontinuities present in filters which were deployed in the work. This approach is compared with the proposed work in Section 4.

The edge gradient (defined as edginess) is computed using 1-D processing of image which does not use the filter with discontinuities [20]. Here, smoothing operator is applied along one direction and it's derivative operator is applied along the orthogonal direction. The same procedure is followed along the orthogonal direction. Combination of these two partial edge gradients can give intensity gradient (direction as well as magnitude) of the image. Let \mathbf{x}_0 denotes the vertical edge evidence of image \mathbf{x} obtained after applying \mathbf{E}_0 operator (derivative operator along each row of image). Similarly, the horizontal edge evidence \mathbf{x}_{90} can be obtained by applying \mathbf{E}_{90} operator (derivative operator along each column of image), and the magnitude of edge gradient (\mathbf{x}_g) as:

$$\mathbf{x}_g = \sqrt{\mathbf{x}_0^2 + \mathbf{x}_{90}^2}. \quad (4)$$

Let this operation of the equation is denoted by operator \mathbf{E}_g . Fig. 1 shows the edge evidences obtained for an image. One can observe from the figure that edge contains perceptually significant information of the image.

In the proposed work, a constraint is added in equation (3),

where the difference in edge information of the reconstructed (down-sampled version) HR image and LR image is minimized. Let \mathbf{e}_d denotes the difference in edge evidences, derived using operator $\mathbf{D} \in \{\mathbf{E}_0, \mathbf{E}_{90}, \mathbf{E}_g\}$, of LR image and down-sampled reconstructed HR image. Thus equation (3) become:

$$\begin{aligned} \hat{\mathbf{c}} &= \arg \min_{\mathbf{c}} \{\|\mathbf{y} - \mathbf{SBAc}\|_2^2 + \lambda \|\mathbf{c}\|_1 + \beta \|\mathbf{e}_d\|_2^2\} \\ &= \arg \min_{\mathbf{c}} \{\|\mathbf{y} - \mathbf{SBAc}\|_2^2 + \lambda \|\mathbf{c}\|_1 \\ &\quad + \beta \|\mathbf{D}\{\mathbf{y}\} - \mathbf{D}\{\mathbf{SBAc}\}\|_2^2\}. \end{aligned} \quad (5)$$

Thus, the resultant sparse coefficients will be closer to the sparse coefficients of the unknown original HR image. In equation (5), the parameters λ and β are very important because they assign weights to various regularization constraints. The parameters are computed using MAP (Maximum A Posteriori) estimation [8], which assumes that the difference in edge evidences \mathbf{e}_d follows Laplacian model and β is computed by maximizing the following equation

$$\hat{\mathbf{e}}_d = \arg \max_{\mathbf{e}_d} \{\log P(\mathbf{e}_d | \mathbf{y})\}. \quad (6)$$

The resultant β will be

$$\beta = \frac{2\sqrt{2}\sigma_n^2}{\sigma_i + \epsilon}, \quad (7)$$

where σ_i is standard deviation of the set $e_{d,i}$ (i -th element of \mathbf{e}_d), ϵ is a small constant and σ_n is standard deviation of noise (assuming that the degraded image is contaminated with additive white Gaussian noise). Similar relation can be derived for λ . Now the equation (5) can be solved using iterative shrinkage algorithm as explained in [22]. The proposed algorithm is summarized in Table 1.

Figs. 2 (d), (e) and (f) show the SR images derived for a synthetic image (Fig. 2 (a)) using \mathbf{E}_0 , \mathbf{E}_{90} and \mathbf{E}_g operator in edge preserving constraints of equation (5), respectively. One can observe that the vertical edges are less smeared in Fig. 2 (d) as compared to Figs. 2 (e) and (f) because the proposed approach attempts to preserves the vertical edge evidences. Similar results can be observed from Fig. 2 (e) also.

4. EXPERIMENTAL RESULTS

The performance of the proposed approach is compared along with existing approaches in Table 2 using HR images mentioned in [23, 16, 4, 8]. Each HR image (name given in first column of Table 2), is of size 256×256 , is blurred using 7×7 Gaussian kernel of zero mean and standard deviation 1.6. The blurred images are down-sampled by scale factor 3 in both the directions (horizontal and vertical) to generate the LR images, which need to be super-resolved.

Table 1. Steps of the proposed approach

Task: To find HR image \mathbf{x} from the LR image \mathbf{y} .

Available Data: LR image (\mathbf{y}), scale-up factor (s), the trained dictionary (\mathbf{A}) and the blur kernel (\mathbf{B}).

Initialization: Initialize $k = 0$ and

- Set the stopping criteria: 1) Maximum iteration number = M and 2) Error Threshold = Th .
- Set the initial parameters λ, β equal to 1.
- Upscale the input LR image by factor s using bi-cubic interpolation method to get the initial HR approximation of the image.
- Select \mathbf{D} from the set $\{\mathbf{E}_0, \mathbf{E}_{90}, \mathbf{E}_g\}$.

Iteration: Set $k = k + 1$ and apply the following steps:

- Slice the estimated HR image into patches of same dimension as \mathbf{A} is derived.
- Select the dictionary adaptively for each patch.
- Solve the optimization equation (5) to get $\hat{\mathbf{c}}_j$ using Iterative Shrinkage algorithm [22].
- Update λ and β following the same method as in equation (7).
- Get the full image from the coefficients and the dictionary $\hat{\mathbf{x}} = \mathbf{A} \mathbf{c} = (\sum_{j=1}^N \mathbf{P}_j^T \mathbf{P}_j)^{-1} \sum_{j=1}^N (\mathbf{P}_j^T \mathbf{A} \mathbf{c}_j)$, where \mathbf{P}_j is a matrix that is responsible for patch extraction \mathbf{x}_j from the image \mathbf{x} , i.e. $\mathbf{x}_j = \mathbf{P}_j \mathbf{x}$ and $j = 1, 2, \dots, N$ (No. of patches).

Stopping Rule: If $\|\mathbf{x}_k - \mathbf{x}_{k-1}\|_2^2 \leq Th$, stop. Otherwise iterate till $k = M$.

Output: The result is HR image \mathbf{x} .

Table 2. Results

Images	TV [23]		Raw Patch [16]		Steering Kernel [4]		ASDS-AR-NL [8]		Edge Preserving(\mathbf{E}_0)		Edge Preserving(\mathbf{E}_{90})		Edge Preserving(\mathbf{E}_g)	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bike	23.61	0.7567	23.20	0.7188	24.38	0.7960	24.62	0.7962	24.78	0.8045	24.79	0.8045	24.77	0.8039
Butterfly	26.60	0.9036	23.73	0.7942	26.93	0.9090	27.34	0.9047	28.70	0.9264	28.67	0.9259	28.70	0.9265
Flower	27.38	0.8111	27.76	0.7929	28.86	0.8460	29.19	0.8480	29.53	0.8576	29.56	0.8582	29.56	0.8583
Girl	31.21	0.7878	32.51	0.7912	33.44	0.8230	33.53	0.8242	34.89	0.8624	34.90	0.8626	34.90	0.8625
Hat	29.19	0.8569	29.65	0.8362	30.81	0.8750	30.93	0.8706	31.40	0.8733	31.39	0.8732	31.41	0.8734
Parrot	27.59	0.8856	27.98	0.8665	29.93	0.9110	30.00	0.9093	30.60	0.9166	30.60	0.9165	30.59	0.9166
Plants	31.28	0.8784	31.48	0.8698	33.27	0.9130	33.47	0.9095	34.11	0.9183	34.09	0.9183	34.06	0.9181
Average	28.12	0.8400	28.04	0.8099	29.66	0.8676	29.87	0.8661	30.57	0.8799	30.57	0.8799	30.57	0.8799

Note: The bold fonts represent best values in that row.

For color images the luminance component and chromatic component are separated, and luminance component is processed to super-resolve because human eyes are relatively more sensitive to luminance changes. Chromatic component is scaled up using bi-cubic interpolation method and later both upscaled components are combined to get final super-resolved image.

We have used two objective parameters namely: structural similarity index measure (SSIM) and peak-signal-to-noise ratio (PSNR) to measure the perceptual quality of SR images [24]. The SSIM (varies between 0 and 1) is considered to be correlated with the perceptual quality of the human visual system (HVS). If SSIM is closer to 1 then the SR image is closer to the original image. Similarly higher PSNR values represent better quality of images.

In proposed approach and ASDS-AR-NL [8] approach, the dictionary is learned from several HR images as explained in section 2. Experimentally we have found that dictionary with patch size of 5×5 produces the best results. In Table 2, we have also shown the results when different edge evidences (horizontal, vertical and magnitude of edge) are used separately in the proposed approach. One can observe that the proposed approach performs better than the existing approaches, which emphasizes the point that the edge plays significant role in perceptual quality of image and has to be preserved in SR process. The results also show that proposed approach performs better than the approach given in [16], which proves that not only edge preservation is important but also how are they computed. The objective parameters for SR images derived using $\mathbf{E}_0, \mathbf{E}_{90}$ and \mathbf{E}_g in the proposed approach are more or less similar. This is due to fact that all the images used in the Table 2 are natural images and does not have specific horizontal or vertical edges. But perceptually, differences can be observed in Fig. 3, which shows the results of

proposed approach using one of the images mentioned in Table 2.

5. SUMMARY

This paper explains the importance of edge preserving constraints in sparse coding based SR image generation. Three different constraints are used which attempt to preserve the horizontal, vertical and magnitude of edge separately. The edge information is derived using 1-D processing of the image. The experimental results show the smearing of edges in the derived HR image is less in the proposed approach as compared to the existing approaches. The results can be further improved by combining the HR results obtained from different edge preserving constraints.

6. REFERENCES

- [1] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang, "Super-resolution image reconstruction: a technical overview," *IEEE, Signal Processing Magazine*, vol. 20, no. 3, pp. 21 – 36, may 2003.
- [2] Qiangqiang Yuan, Liangpei Zhang, and Huanfeng Shen, "Multiframe super-resolution employing a spatially weighted total variation model," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 3, pp. 379 – 392, march 2012.
- [3] W.T. Freeman, T.R. Jones, and E.C. Pasztor, "Example-based super-resolution," *IEEE, Computer Graphics and Applications*, vol. 22, no. 2, pp. 56 – 65, mar/apr 2002.

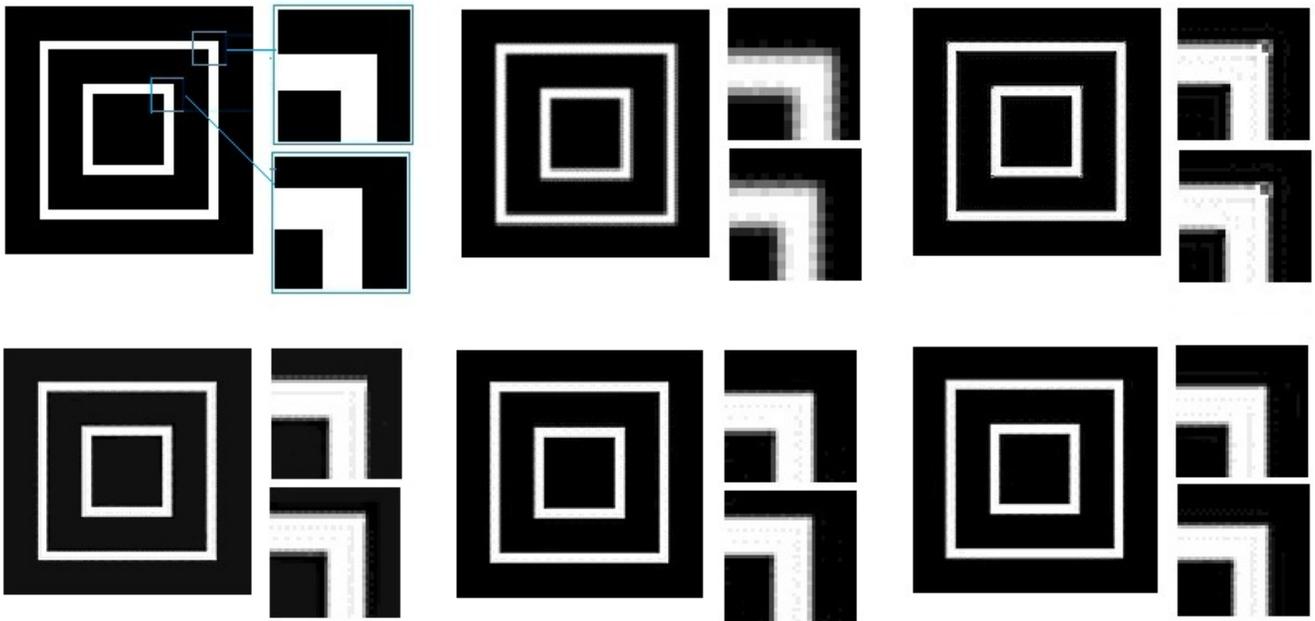


Fig. 2. SR results of *Artificial image* Top left (a): Original image, Top middle (b): LR image, Top Right (c): ASDS-AR-NL result, Bottom left (d): Proposed approach result (vertical edge preservation), Bottom middle (e): Proposed approach result (horizontal edge preservation), Bottom right (f): Proposed approach result (edge gradient magnitude preservation).

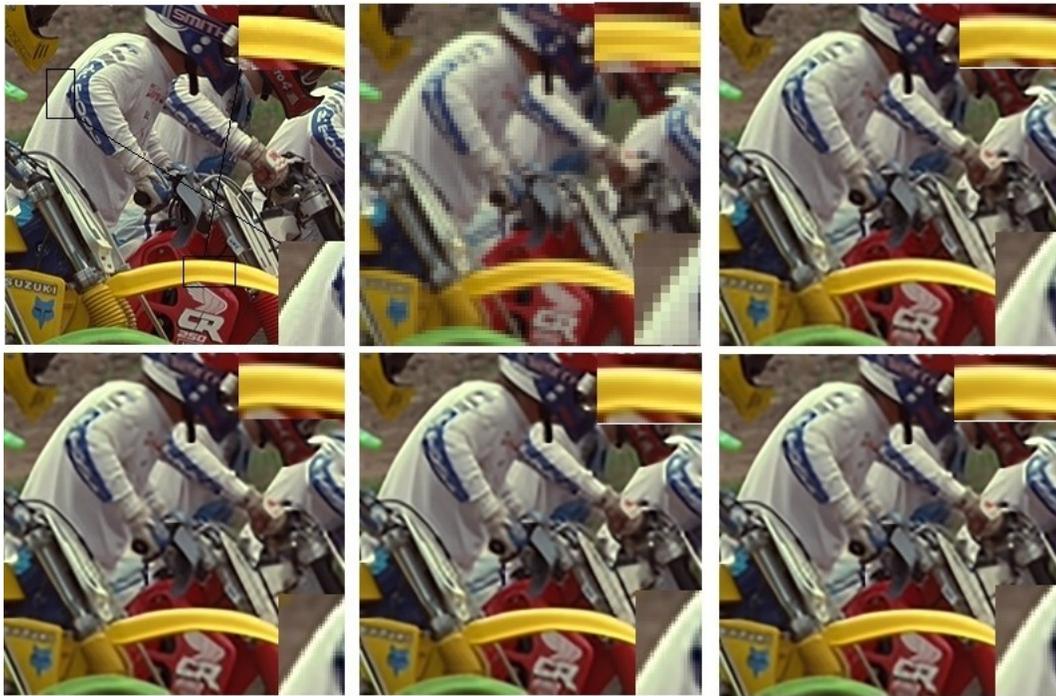


Fig. 3. SR results of *Bike image* Top left (a): Original image, Top middle (b): LR image, Top Right (c): ASDS-AR-NL result, Bottom left (d): Proposed approach result (vertical edge preservation), Bottom middle (e): Proposed approach result (horizontal edge preservation), Bottom right (f): Proposed approach result (edge gradient magnitude preservation).

- [4] Kaibing Zhang, Xinbo Gao, Dacheng Tao, and Xuelong Li, "Single image super-resolution with non-local means and steering kernel regression," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4544–4556, nov. 2012.
- [5] Xinbo Gao, Kaibing Zhang, Dacheng Tao, and Xuelong Li, "Joint learning for single-image super-resolution via a coupled constraint," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 469–480, 2012.
- [6] Xinbo Gao, Kaibing Zhang, Dacheng Tao, and Xuelong Li, "Image super-resolution with sparse neighbor embedding," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3194–3205, 2012.
- [7] Kaibing Zhang, Xinbo Gao, Dacheng Tao, and Xuelong Li, "Multi-scale dictionary for single image super-resolution," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012*, 2012, pp. 1114–1121.
- [8] Weisheng Dong, Lei Zhang, Guangming Shi, and Xiaolin Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, jul 2011.
- [9] N.P. Galatsanos and A.K. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *IEEE Transactions on Image Processing*, vol. 1, no. 3, pp. 322–336, jul 1992.
- [10] Leonid I. Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 14, pp. 259–268, 1992.
- [11] J.M. Bioucas-Dias, "Bayesian wavelet-based image deconvolution: a gem algorithm exploiting a class of heavy-tailed priors," *IEEE Transactions on Image Processing*, vol. 15, no. 4, pp. 937–951, april 2006.
- [12] D.L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, april 2006.
- [13] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, dec. 2006.
- [14] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *IEEE 12th International Conference on Computer Vision*, 29 2009-oct. 2 2009, pp. 2272–2279.
- [15] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on Image Processing*, vol. 17, no. 1, pp. 53–69, jan. 2008.
- [16] Jianchao Yang, J. Wright, T. Huang, and Yi Ma, "Image super-resolution as sparse representation of raw image patches," in *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008.*, june 2008, pp. 1–8.
- [17] M. Elad, M.A.T. Figueiredo, and Yi Ma, "On the role of sparse and redundant representations in image processing," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 972–982, june 2010.
- [18] David L. Donoho, "For most large underdetermined systems of equations, the minimal l_1 -norm near-solution approximates the sparsest near-solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 7, pp. 907–934, 2006.
- [19] Weisheng Dong, Lei Zhang, and Guangming Shi, "Centralized sparse representation for image restoration," in *2011 IEEE International Conference on Computer Vision (ICCV)*, nov. 2011, pp. 1259–1266.
- [20] Anil Kumar Sao, B. Yegnanarayana, and B.V.K. Vijaya Kumar, "Significance of image representation for face verification," *Signal, Image and Video Processing*, vol. 1, pp. 225–237, 2007.
- [21] R. Rubinstein, A.M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, june 2010.
- [22] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [23] Antonio Marquina and Stanley J. Osher, "Image super-resolution by TV-regularization and bregman iteration," *Journal of Scientific Computing*, vol. 37, pp. 367–382, 2008.
- [24] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, aug. 2010, pp. 2366–2369.