

ROBUST SUPER-RESOLUTION USING MULTIPLE BASES AND 3D FILTERING

Naushad Ansari, Lin Weisi

Nanyang Technological University, Singapore

ABSTRACT

In this work, we propose a method for the super-resolution of images in the presence of impulse noise. First, the impulse noise locations are identified using a detector and then, an optimization problem is solved to reconstruct the high-resolution image. Further, we propose the concept of image reconstruction using multiple bases and 3D filtering to improve the performance of the proposed method. We call this concept as MB3D. We apply the proposed method along with MB3D on various datasets to test its efficacy. We also perform experiments on some real noisy images.

Index Terms— Super-resolution, sparse recovery, impulse noise.

1. INTRODUCTION

There is a requirement of high-resolution images in many areas such as multimedia, surveillance, remote-sensing etc. However, limitations of solid state sensors restrict the spatial resolution of captured images [1]. This triggers the need for super-resolution (SR) algorithms, which produce a high resolution (HR) image from one or more low resolution (LR) images captured. Actually, there exist many applications where one cannot have more than one LR image such as the restoration of old photos, calligraphy, or paintings etc. In such applications, single image super-resolution (SISR) has its own importance. Hence, we only focus on SISR in this work.

Methods for SISR can be broadly classified into three categories: interpolation based methods, example based methods, and deep learning based methods. Interpolation based methods [2, 3, 4] are computationally efficient but they tend to produce smooth images with ringing and blurring artifacts. Example based methods [5, 6, 7, 8, 9, 10] learn a relationship between LR and HR patches of the training dataset assuming that LR and HR patches have similar local geometry. This relationship is used to obtain the HR patch from a given LR patch. Deep learning based methods [11, 12, 13] learn the mapping between training HR images and the upscaled version of training LR images to obtain a high-resolution image from a given LR image.

Although SISR has achieved great success, existing SR algorithms often face challenge of being sensitive to noise while dealing with degraded images. Impulse noise is a com-

mon cause of image degradation due to either faulty camera sensors, faulty memory locations, or timing errors in analog to digital converters [14]. There exist some work to perform SR in the presence of impulse noise [15, 16]. However, these methods are for multi-image super-resolution. For SISR, authors in [17] proposed to use l_1 fidelity with l_1 regularization.

It is observed in the literature that multi-stage methods perform better than l_1 fidelity based methods for impulse denoising (not SR) of images [18]. Hence, we propose to use a two-stage method for SR in the presence of impulse noise in this paper. In the first stage, faulty/corrupted pixel locations are identified using a detector and in the second stage, an optimization problem is solved for SR using wavelet as the sparsifying transform and using the information of faulty locations from the first stage.

There exist multiple wavelet bases and no wavelet basis is optimal (in terms of reconstruction) for all images. One wavelet basis may be optimal for one image while the same basis is not optimal for another image. We got the same observation even at pixel level. One pixel is best reconstructed with one wavelet while other pixels are best reconstructed with some other wavelets. This motivated us to use multiple wavelet bases simultaneously and use the reconstruction ability of all bases used. This led us to propose the concept of image reconstruction with multiple wavelet bases followed by 3D filtering in this paper. We call this concept as MB3D. We observe that MB3D is able to improve the performance of all wavelets used.

Salient contributions of this works are as follows: 1) We propose a method for SISR in the presence of impulse noise. 2) We propose the concept of MB3D to further improve the performance of the proposed method.

2. PROPOSED METHOD

2.1. SISR in the presence of impulse noise

SISR can be mathematically modeled as below:

$$\mathbf{x}_l = \mathbf{D}\mathbf{H}\mathbf{x}_h, \quad (1)$$

where \mathbf{x}_l and \mathbf{x}_h are the vectorized form of the LR and HR image, respectively. \mathbf{D} and \mathbf{H} are the downsampling and blurring operator, respectively.

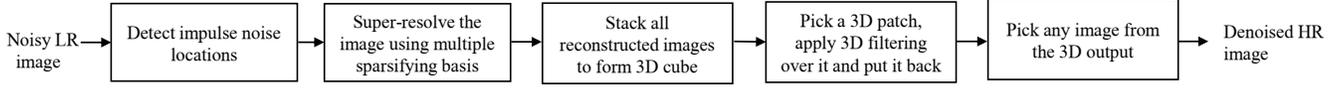


Fig. 1: Block diagram of the proposed method

If the image is corrupted by impulse noise, we have the following noisy LR image:

$$\mathbf{x}_n = \mathcal{N}(\mathbf{x}_l), \quad (2)$$

where $\mathcal{N}(\cdot)$ corrupts a pixel with probability p .

For the reconstruction of the HR image, we follow a two-stage method. In the first stage, we identify the impulse noise locations using the adaptive center-weighted median filter (ACWMF) [19]. Let Ω be the operator that picks non-noisy (identified as non-corrupted by ACWMF) pixels from \mathbf{x}_n and store them in \mathbf{y} as: $\mathbf{y} = \Omega(\mathbf{x}_n)$. In the second stage, we solve the following optimization problem to obtain the HR image:

$$\underset{\mathbf{x}_h}{\text{minimize}} \|\mathbf{y} - \Omega(\mathbf{D}\mathbf{H}\mathbf{x}_h)\|_p^p + \lambda \|\mathcal{W}\mathbf{x}_h\|_q^q. \quad (3)$$

Here, $0 \leq q \leq 1$. If no detector is used, $p = 2$ for Gaussian denoising and $p = 1$ for impulse denoising can be used. In this way, the above formulation is in very general form and can be used in various cases. Also, \mathcal{W} is the wavelet transform and λ is the regularization parameter. Writing $\Omega(\mathbf{D}\mathbf{H}\mathbf{x}_h) = \mathbf{A}\mathbf{x}_h$ in the above formulation and removing subscript from \mathbf{x}_h , we obtain the following optimization problem:

$$\underset{\mathbf{x}}{\text{minimize}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_p^p + \lambda \|\mathcal{W}\mathbf{x}\|_q^q. \quad (4)$$

As the above problem cannot be solved directly, we solve the above problem using split-Bregman approach in the following subsection in the context of MB3D.

2.2. Reconstruction improvement using MB3D

As discussed in section 1, different wavelets reconstruct a pixel with different accuracy. Hence, we use multiple wavelet bases to reconstruct an image in order to use the reconstruction ability of all wavelets used. This step is followed by 3D filtering. The whole procedure is described below.

Using multiple wavelet bases, we solve the following optimization problem:

$$\underset{\mathbf{X}}{\text{minimize}} \sum_{i=1}^{N_b} \|\mathbf{y} - \mathbf{A}\mathbf{x}_i\|_p^p + \lambda \|\mathcal{W}_i\mathbf{x}_i\|_q^q. \quad (5)$$

Here, N_b is the number of wavelet bases used in the reconstruction process and \mathcal{W}_i refers to the i^{th} wavelet basis used to reconstructs image \mathbf{x}_i . $\mathbf{X} = \mathcal{F}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_b})$ and $\mathcal{F}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_b}) = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_{N_b}]$, i.e., the operator \mathcal{F} stacks images reconstructed with various wavelets as columns in the matrix \mathbf{X} .

We solve the above formulation using split-Bregman approach [20]. Introduce proxy variables, $\mathbf{u}_i = \mathbf{y} - \mathbf{A}\mathbf{x}_i$ and $\mathbf{v}_i = \mathcal{W}_i\mathbf{x}_i$. The resultant constrained constrained problem can be converted to the following unconstrained optimization problem using Bregman variables \mathbf{b}_i and \mathbf{c}_i :

$$\underset{\mathbf{X}, \mathbf{U}, \mathbf{V}}{\text{minimize}} \sum_{i=1}^{N_b} \{ \|\mathbf{u}_i\|_p^p + \lambda \|\mathbf{v}_i\|_q^q + \mu_1 \|\mathbf{u}_i - \mathbf{y} + \mathbf{A}\mathbf{x}_i - \mathbf{b}_i\|_2^2 + \mu_2 \|\mathbf{v}_i - \mathcal{W}_i\mathbf{x}_i - \mathbf{c}_i\|_2^2 \},$$

where $\mathbf{U} = \mathcal{F}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{N_b})$ and $\mathbf{V} = \mathcal{F}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N_b})$.

The above equation can be seen as following three sub-problems which can be solved independently:

$$\min_{\mathbf{X}} \sum_{i=1}^{N_b} \mu_1 \|\mathbf{u}_i - \mathbf{y} + \mathbf{A}\mathbf{x}_i - \mathbf{b}_i\|_2^2 + \mu_2 \|\mathbf{v}_i - \mathcal{W}_i\mathbf{x}_i - \mathbf{c}_i\|_2^2 \quad (6)$$

$$\underset{\mathbf{U}}{\text{minimize}} \sum_{i=1}^{N_b} \|\mathbf{u}_i\|_p^p + \mu_1 \|\mathbf{u}_i - \mathbf{y} + \mathbf{A}\mathbf{x}_i - \mathbf{b}_i\|_2^2 \quad (7)$$

$$\underset{\mathbf{V}}{\text{minimize}} \sum_{i=1}^{N_b} \lambda \|\mathbf{v}_i\|_q^q + \mu_2 \|\mathbf{v}_i - \mathcal{W}_i\mathbf{x}_i - \mathbf{c}_i\|_2^2 \quad (8)$$

Equation (6) has the following closed form solution:

$$\hat{\mathbf{x}}_i = (\mu_1 \mathbf{A}'\mathbf{A} + \mu_2 \mathcal{W}_i' \mathcal{W}_i)^{-1} (\mu_2 \mathcal{W}_i' \mathbf{y}_{1i} - \mu_1 \mathbf{A}' \mathbf{y}_{2i}), \quad \forall i \in \{1, 2, \dots, N_b\} \quad (9)$$

where $\mathbf{y}_{1i} = \mathbf{v}_i - \mathbf{c}_i$ and $\mathbf{y}_{2i} = \mathbf{u}_i - \mathbf{y} - \mathbf{b}_i$. Equations (7) and (8) are l_p regularized least squares which can be solved using iterative p-shrinkage (IPS) [21] as below:

$$\begin{aligned} \hat{\mathbf{u}}_i &= \mathcal{S}(\mathbf{y} - \mathbf{A}\mathbf{x}_i + \mathbf{b}_i, \frac{1}{2\mu_1}, p), \quad \forall i \in \{1, 2, \dots, N_b\} \\ \hat{\mathbf{v}}_i &= \mathcal{S}(\mathcal{W}_i\mathbf{x}_i + \mathbf{c}_i, \frac{\lambda}{2\mu_2}, q), \quad \forall i \in \{1, 2, \dots, N_b\}, \end{aligned} \quad (10)$$

where, \mathcal{S} is the following function:

$$\mathcal{S}(\mathbf{x}, t, p) = \max\{|\mathbf{x}| - t^{2-p}|\mathbf{x}|^{p-1}, 0\} \odot \text{sign}(\mathbf{x}), \quad (11)$$

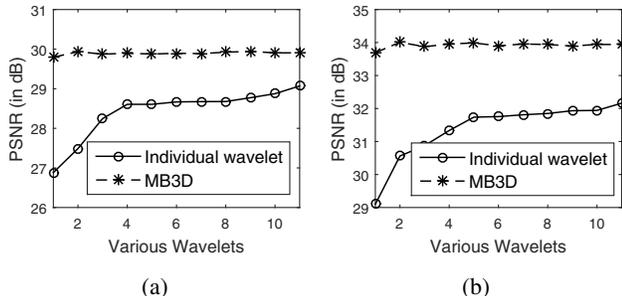
where t is the threshold, $|\cdot|$ denotes the absolute value of the vector, and \odot denotes the point-wise multiplication operator.

At the end, Bregman variables are updated using the following relation:

$$\begin{aligned} \mathbf{b}_i &= \mathbf{b}_i - \mathbf{u}_i + \mathbf{y} - \mathbf{A}\mathbf{x}_i, \quad \forall i \in \{1, 2, \dots, N_b\} \\ \mathbf{c}_i &= \mathbf{c}_i - \mathbf{v}_i + \mathcal{W}_i\mathbf{x}_i, \quad \forall i \in \{1, 2, \dots, N_b\} \end{aligned} \quad (12)$$

Table 1: Performance of various methods on various datasets in terms of PSNR (in dB)

Method	Average PSNR in dB over different datasets														
	Set5					Set10					Set14				
	5	10	15	20	25	5	10	15	20	25	5	10	15	20	25
$l_1 - l_1$ [17]	28.0	27.3	26.5	25.8	24.5	28.0	27.4	26.9	26.2	25.4	25.8	25.3	24.7	24.1	23.3
Med-wdsr	27.2	26.5	25.5	24.2	22.9	27.0	26.4	25.4	24.3	23.4	25.0	24.4	23.8	23.0	22.1
l_0 TV-bic	28.2	26.8	26.6	25.8	24.7	28.2	27.4	26.6	25.9	25.4	25.4	24.6	24.1	23.6	23.0
l_0 TV-wdsr	27.9	26.1	25.6	24.7	23.5	27.9	26.8	25.7	24.8	24.4	25.5	24.4	23.6	23.0	22.3
Proposed	29.3	28.6	27.7	27.0	26.1	28.9	28.2	27.6	26.8	25.8	26.0	25.6	25.0	24.4	23.8
Proposed-ID	33.0	32.5	31.9	31.2	30.3	31.6	31.3	31.0	30.5	30.0	29.4	29.1	28.7	28.1	27.6

**Fig. 2:** Reconstruction results (in terms of PSNR in dB) of image (a) ‘Lena’ with 15% impulse noise, (b) ‘House’ with 20% impulse noise. Up-sampling factor=2.

The above procedure is repeated for either a fixed number of iterations or until the solution converges. We use 50 iterations in our experiments as the solution was observed to converged before that.

Once we obtain reconstructed images with all bases in the form of a 3D cube, i.e., matrix \mathbf{X} , we extract 3D patches from this cube and apply 3-D filtering to it. 3D filtering consists of two steps. In the first step, we compute the 3D DCT transform of the patch, apply soft thresholding to the resultant patch, and compute the inverse 3D transform of it. In the second step, we vectorize each 2D plane of the 3D patch to obtain a matrix and approximate the resultant matrix to rank-1 matrix. In the end, we reshape the matrix to obtain the 3D patch and put it back to its location.

In our experiment, we use the patch size of 8×8 and we use the threshold of 0.04 for soft-thresholding of the patch. The complete block diagram of the proposed method is shown in Fig. 1.

3. EXPERIMENT

In this section, we present various experiments to show the efficacy of the proposed method. We perform all experiments on random-valued impulse noise. We evaluate the performance in terms of peak signal to noise ratio (PSNR) in dB. We used $\lambda = 1$, $\mu_1 = 10$, $\mu_2 = 10$, $p = 2$, and $q = 1$ in all our experiments. We used 11 different orthogonal and bi-orthogonal wavelets: ‘db1’, ‘db2’, ‘db4’, ‘db10’, ‘bior2.2’, ‘bior4.4’, ‘sym6’, ‘sym8’, ‘sym10’, ‘coif4’ and ‘coif5’. Please note that these wavelets show different performance on patches depending upon which wavelet is most correlated with that patch.

3.1. Effect of MB3D on robust SR

To show the effect of MB3D, we perform SR with an up-sampling factor of 2 on images ‘Lena’ and ‘House’ with 15% and 20% random-valued impulse noise, respectively. Results are shown in Fig. 2a and Fig. 2b, respectively. The x-axis represents various wavelets with their sorted performance in terms of PSNR (shown on the y-axis). It can be observed that: 1) Reconstruction performance varies too much with wavelets when individual wavelets are used without MB3D. Specifically, there is a performance difference of 2 dB and 3 dB in PSNR between poorly performing and best-performing wavelets while reconstructing image ‘Lena’ and ‘House’, respectively. 2) MB3D not only equalizes the performance of every wavelet, it enhances the performance over best performing wavelet. 3) There is a performance improvement of 3 dB and 4 dB over the poorly performing wavelet, while there is an improvement of 0.9 dB and 1.8 dB over best-performing wavelet. As reconstruction results are similar with all wavelets after MB3D, one can choose output of any wavelet after MB3D. This shows the ability of MB3D to enhance the reconstruction performance over various wavelets.

3.2. Comparison with existing methods

In this sub-section, we compare the performance of the proposed method with several of its competitors. We perform experiments on three datasets. Two of these datasets are widely used for SR algorithms. These are Set5 and Set14. We also perform experiments on the dataset used in [22]. This dataset contains 10 images of different frequency characteristics, 3 images are rich in low frequency, 3 images are rich in high frequency, and rest have varied low and high-frequency content. We use this dataset as it makes more sense to apply an algorithm on images of variable frequency content as performance varies over different frequency content images. We call this dataset as ‘Set10’. We perform experiments with 5% to 25% random values impulse noise.

We compare the performance with existing $l_1 - l_1$ method [17]. As impulse denoising followed by SR is one category of methods for robust SR, we consider three such methods in this category: Med-wdsr, l_0 TV-bic, and l_0 TV-wdsr. Med-wdsr stands for median filtering (simple and fast method for impulse denoising) followed by wdsr [23] for SR. l_0 TV-bic

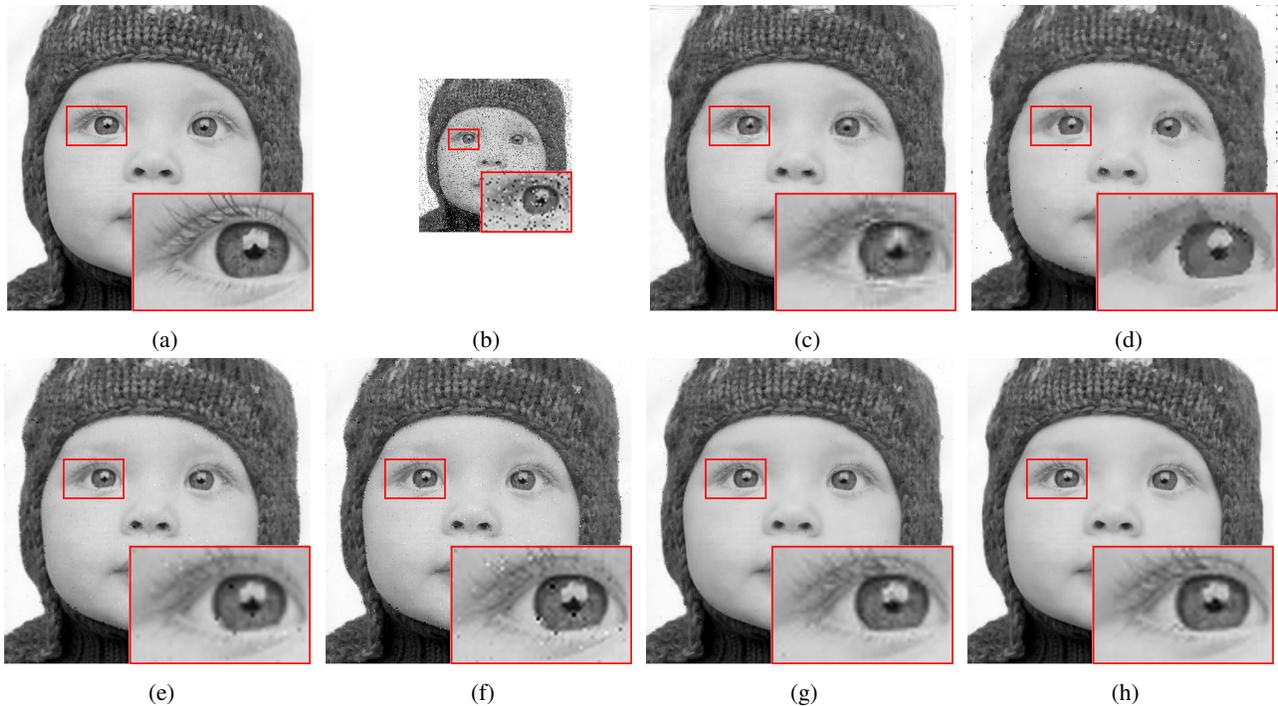


Fig. 3: Visual comparison of various methods on one image of ‘Set5’ dataset at 15% impulse noise and up-sampling factor=2. (a): Original high-resolution image, (b) noisy low-resolution image; image reconstructed with (c) $l_1 - l_1$, (d) Med-wdsr, (e) l_0 TV-bic, (f) l_0 TV-wdsr, (g) Proposed method, and (h) Proposed method with ideal detector (Proposed-ID)

and l_0 TV-wdsr stands for l_0 TV [24] for impulse denoising followed by bicubic interpolation and wdsr for SR, respectively. Please note that wdsr and l_0 TV are state-of-the-art methods for SR and impulse denoising, respectively. Table 1 shows all these results. We also use the proposed method with the ideal detector, shown as Proposed-ID (performance in blue color) in the table. This uses the actual locations of the faulty pixels instead of using ACWMF as the detector. We also tried wdsr alone or SR followed by impulse denoising but results were very poor with them.

It can be observed from the table that the proposed method is better than all other methods. Also, the performance with the Proposed-ID is much better than the proposed method which shows the capability of the method to reconstruct images with much higher accuracy if an ideal detector is present.

Visual results on one of the images of ‘Set5’ dataset at 15% noise ratio and up-sampling factor=2 is shown in Fig. 3. Eye of the ‘baby’ is zoomed to have a better visualization of results. We can see that $l_1 - l_1$ has several artifacts while Med-wdsr produces very smooth images with very fewer details preserved. l_0 TV-bic and l_0 TV-wdsr has noisy samples remaining. The images reconstructed with Proposed and Proposed-ID are very close and free from such artifacts.

3.3. Experiment on real noisy images

We show the application of the proposed method on scratched images (Fig. 4a and 4c) as scratches can be viewed as impulse noise in the image [24]. Fig. 4b and 4d show SR images

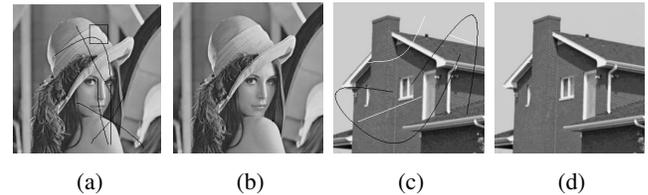


Fig. 4: Visual results of SR on real scratched low resolution images with up-sampling factor=2 using the proposed method. (a) and (c) low resolution scratched images; (b) and (d) reconstructed HR image.

reconstructed with up-sampling factor=2 using the proposed method.

4. CONCLUSION AND FUTURE WORK

A method for super-resolution in the presence of impulse noise has been proposed. The concept of image reconstruction using multiple bases and 3D filtering (MB3D) has also been proposed to further improve the performance. Although here MB3D has been applied in the application of robust super-resolution only, it can be further explored in several inverse problems such as denoising, compressed sensing, deblurring etc.

Acknowledgment

I would like to thank Dr. Anubha Gupta and Dr. Hemant Kumar Aggarwal for their valuable suggestions and discussions.

5. REFERENCES

- [1] K. Nasrollahi and T. B. Moeslund, "Super-resolution: a comprehensive survey," *Machine vision and applications*, vol. 25, no. 6, pp. 1423–1468, 2014.
- [2] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE transactions on acoustics, speech, and signal processing*, vol. 29, no. 6, pp. 1153–1160, 1981.
- [3] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE transactions on image processing*, vol. 10, no. 10, pp. 1521–1527, 2001.
- [4] C. E. Duchon, "Lanczos filtering in one and two dimensions," *Journal of applied meteorology*, vol. 18, no. 8, pp. 1016–1022, 1979.
- [5] 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, 2002.
- [6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE transactions on image processing*, vol. 19, no. 11, pp. 2861–2873, 2010.
- [7] Z. Xiong, D. Xu, X. Sun, and F. Wu, "Example-based super-resolution with soft information and decision," *IEEE Transactions on Multimedia*, vol. 15, no. 6, pp. 1458–1465, 2013.
- [8] K. Zhang, D. Tao, X. Gao, X. Li, and Z. Xiong, "Learning multiple linear mappings for efficient single image super-resolution," *IEEE Transactions on Image Processing*, vol. 24, no. 3, pp. 846–861, 2015.
- [9] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1. IEEE, 2004, pp. I–I.
- [10] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5197–5206.
- [11] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2016.
- [12] J. Kim, J. Kwon Lee, and K. Mu Lee, "Deeply-recursive convolutional network for image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1637–1645.
- [13] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang, "Deep laplacian pyramid networks for fast and accurate superresolution," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, no. 3, 2017, p. 5.
- [14] R. C. Gonzalez, R. E. Woods *et al.*, "Digital image processing," 2002.
- [15] L. Yue, H. Shen, Q. Yuan, and L. Zhang, "A locally adaptive l1- l2 norm for multi-frame super-resolution of images with mixed noise and outliers," *Signal Processing*, vol. 105, pp. 156–174, 2014.
- [16] H. Shen, L. Peng, L. Yue, Q. Yuan, and L. Zhang, "Adaptive norm selection for regularized image restoration and super-resolution," *IEEE transactions on cybernetics*, vol. 46, no. 6, pp. 1388–1399, 2016.
- [17] J. Jiang, Z. Wang, C. Chen, and T. Lu, "L1-l1 norms for face super-resolution with mixed gaussian-impulse noise," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 2089–2093.
- [18] G. Gao, Y. Liu, and D. Labate, "A two-stage shearlet-based approach for the removal of random-valued impulse noise in images," *Journal of Visual Communication and Image Representation*, vol. 32, pp. 83–94, 2015.
- [19] T.-C. Lin, "A new adaptive center weighted median filter for suppressing impulsive noise in images," *Information Sciences*, vol. 177, no. 4, pp. 1073–1087, 2007.
- [20] T. Goldstein and S. Osher, "The split bregman method for l1-regularized problems," *SIAM journal on imaging sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [21] J. Woodworth and R. Chartrand, "Compressed sensing recovery via nonconvex shrinkage penalties," *Inverse Problems*, vol. 32, no. 7, p. 075004, 2016.
- [22] N. Ansari and A. Gupta, "Image reconstruction using matched wavelet estimated from data sensed compressively using partial canonical identity matrix," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 3680–3695, 2017.
- [23] C. Cruz, R. Mehta, V. Katkovnik, and K. O. Egiazarian, "Single image super-resolution based on wiener filter in similarity domain," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1376–1389, 2018.
- [24] G. Yuan and B. Ghanem, " $l_{\{0\}}$ tv: A sparse optimization method for impulse noise image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.