Robust Multi-Frame Super-Resolution with Adaptive Norm Choice and Difference Curvature based BTV Regularization

Xiaohong Liu and Jiying Zhao
School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario, Canada, K1N 6N5
Email: xliu151@uottawa.ca, jzhao@uottawa.ca

***Abstract*—Multi-frame super-resolution focuses on reconstructing a high-resolution image from a set of low-resolution images with high similarity. The minimization function derived from maximum a posteriori probability (MAP) is composed of a fidelity term and a regularization term. In this paper, we propose a new fidelity term based on half-quadratic estimation to choose error norm adaptively instead of using fixed L_1 or L_2 norm. Besides, we propose a novel regularization method which combines the advantage of Difference Curvature (DC) and Bilateral Total Variation (BTV) to preserve the edge areas and remove noise simultaneously. The proposed framework is tested on both synthetic data and real data. Our experimental results illustrate the superiority of the proposed method in terms of edge preserving and noise removal over other state-of-the-art algorithms.**

***Index Terms*—Multi-frame super-resolution, difference curvature, half-quadratic estimation, bilateral total variation (BTV)**

I. Introduction

Super-resolution (SR) is a method to increase the image resolution without modifying the sensor of camera. Different from single image super-resolution, multi-frame super-resolution focuses on reconstructing a high-resolution image from a set of low-resolution images with high similarity. It was first addressed in [1] using a frequency domain algorithm which is easy to implement and computationally cheap. But processing multi-frame super-resolution in frequency domain will introduce serious visual artifacts. Since then, many approaches have been proposed to solve the multi-frame SR problem. Because of the limitation of frequency domain approaches, the methods which enhance image in the spatial domain become more and more popular [2, 3]. As super-resolution is an ill-posed problem, regularization techniques are widely used to constrain the minimization function and also regarded as prior knowledge of the related frames. By combining image prior knowledge with fidelity model, Bayesian-based spatial domain methods can effectively solve this ill-posed problem, which makes this kind of methods more popular than others in the field of image super-resolution.

Spatial domain based multi-frame image super-resolution usually reconstructs the high-resolution image from the related low-resolution images by exploiting the subpixel displacements [4]. In practical applications, the subpixel displacements are not only simple affine motion, but also partial movement, non-rigid movement and occlusion. Therefore, the traditional observation models have limited performance to reconstruct high-resolution images [5].

In general, the framework of multi-frame image super-resolution in spatial domain contains two parts. The fidelity term is used to keep the fidelity between the HR frame and LR frames. And the regularization term aims at regularizing the minimization function. Since the noise in observation model usually fits the Gaussian distribution, choosing L_2 norm for fidelity term can obtain good results. But in practical applications, the observation model suffers various noises and errors introduced by inaccurate estimation of registration and blurring kernel. Farsiu *et al.* firstly used L_1 norm rather than L_2 norm in fidelity term and achieved better results than L_2 norm [4]. However, although the L_1 norm is robust for outliers, it may introduce more observation errors than L_2 norm while the estimation of images is accurate. The drawbacks of fixed norms motivated researchers to combine the advantage of L_1 and L_2 norms. Nowadays, some M-estimators such as Huber function [6] were proposed to replace the fixed norms as well. Yue *et al.* [7] proposed a locally adaptive L_1, L_2 norm to handle images with mixed noises and outliers. But by introducing a threshold to choose L_1 or L_2 norm, it makes the minimization function non-derivable. Zeng *et al.* [8] proposed a new method based on half-quadratic estimation to adaptively determine the error norm and the experimental results also illustrate the superiority of their method.

For the regularization techniques, one of the commonly used methods is Tikhonov regularization based on L_2 norm [9]. However, L_2 norm is sensitive to outliers so that it will introduce artifacts into images. Nowadays, sparse prior is very popular in single image super-resolution. But for multi-frame super-resolution, using the redundant information among the low-resolution frames in spatial domain is more reliable than using it in sparse domain. Besides, Total variation (TV) family such as bilateral total variation (BTV) [4] are popular regularization techniques. Farsiu *et al.* showed the BTV could preserve more detail information than Tikhonov regularization and be robust to outliers.

In this paper, we propose a novel robust multi-frame super-resolution method. There are two major contributions which effectively improve the quality of the final estimated HR images:

1) A new fidelity term based on half-quadratic estimation is proposed. In our fidelity term, the half-quadratic estimation is used to choose error norm adaptively according to the change of averaged observation errors rather than employing the traditional fixed L_1 or L_2 norm.
2) A novel Difference Curvature based BTV regularization method (DCBTV) is proposed. Due to the drawbacks of traditional regularization methods, Difference Curvature is adopted to adjust the relevant value in the BTV regularization, which improves the regularized performance in terms of edge preserving and noise removal.

The rest parts of this paper are organized as follows. Section II introduces the observation model and basic framework of multi-frame image super-resolution. Section III details the proposed algorithm which uses half-quadratic estimation for the fidelity term and Difference Curvature for the BTV regularization term. Section IV illustrates the experimental results and Section V concludes this paper.

II. Preliminaries

A. Observation Model of Multi-Frame Super-Resolution

Observation model formulates the relationship between the high-resolution frame and low-resolution frames. In general, low-resolution frames can be regarded as the corresponding high-resolution frame going through the geometric motion operator, blurring operator and down-sampling operator successively. Therefore, the observation model can be formulated as

$$\mathbf{Y}_k = \mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} + \mathbf{n}_k, \tag{1}$$

where $\mathbf{X}$ is the HR frame and expressed in lexicographic order as $\mathbf{X} = [x_1, x_2, ..., x_N]^T$, where N is the total number of pixels in HR frame which equals to $rm \times rn$ and r is the downsampling factor. Therefore, the size of $\mathbf{X}$ is $rm \times rn \times 1$. Similar to the definition of $\mathbf{X}$, $\mathbf{Y}_k = [y_{k,1}, y_{k,2}, ..., y_{k,L}]^T$, which represents the kth LR frame with the size of $mn \times 1$, where $k = 1, 2, ..., K$. K is the number of LR frames and $L = m \times n$. $\mathbf{M}_k$ represents the geometric motion matrix between HR frame and kth LR frame with the size of $rm \times rn \times rm \times rn$. $\mathbf{B}_k$ is the blurring matrix for the kth LR frame with the size of $rm \times rn \times rm \times rn$ and $\mathbf{D}$ is the downsampling matrix with the size of $mn \times rm \times rn$. In general, image noise should be taken into consideration as well. $\mathbf{n}_k$ represents the noise added into the kth LR frame with the size of $mn \times 1$.

B. The Basic Framework of Multi-Frame Super-Resolution

The basic framework of multi-frame super-resolution contains fidelity term and regularization term. For the fidelity term, M-estimator minimizes the residual between the estimated HR frame and given LR frames. The regularization term is used to constrain the minimization function. The traditional framework of multi-frame super-resolution can be formulated as

$$\hat{\mathbf{X}} = arg\min_{\mathbf{X}} \left\{ \sum_{k=1}^{K} \|\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} - \mathbf{Y}_k\|_p^p + \lambda\Upsilon(\mathbf{X}) \right\}, \tag{2}$$

where $\Upsilon(\mathbf{X})$ is the regularization term with respect to $\mathbf{X}$. λ is the trade-off parameter between the two terms and p represents the choise of Lp norm.

For the regularization term $\Upsilon(\mathbf{X})$, image prior knowledge such as Tikhonov regularization and total variation (TV) family are widely used. Equ. (3) shows the expression of the traditional BTV regularization.

$$\Upsilon_{BTV}(\mathbf{X}) = \sum_{l=-P}^{P}\sum_{m=0}^{P} \beta^{|m|+|l|} \left\|\mathbf{X} - \mathbf{S}_x^l\mathbf{S}_y^m\mathbf{X}\right\|_1, \tag{3}$$

where $\mathbf{S}_x^l$ shifts $\mathbf{X}$ by l pixels in horizontal direction and $\mathbf{S}_y^m$ shifts $\mathbf{X}$ by m pixels in vertical direction. β is a scaled weight with the range of $0 < \beta < 1$ and P is a control parameter which controls the decaying effect to the summation of the BTV regularization.

III. Proposed Multi-Frame Super-Resolution Algorithm

In this section, we introduce our proposed algorithm in detail. For the fidelity term, the half-quadratic estimation is used to make norm choice adaptive instead of using fixed L_1 or L_2 norm. For the regularization term, a novel regularization method based on Difference Curvature is proposed to constrain the minimization function.

A. Half-Quadratic Estimation Based Adaptive Fidelity Term

Due to the drawbacks of fixed norms, the half-quadratic function was proposed in [8, 10] to combine the advantage of L_1 and L_2 norms, which is defined as

$$f(x,\alpha) = \alpha\sqrt{\alpha^2 + x^2}, \tag{4}$$

where α is a positive constant. For each LR frame, x represents the observation error which equals to $(\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} - \mathbf{Y}_k)$. The first derivative of $f(x,\alpha)$ with respect to x is shown as:

$$f'(x,\alpha) = \frac{\alpha x}{\sqrt{\alpha^2 + x^2}}. \tag{5}$$

Fig. 1. shows the superiority of half-quadratic function compared with other M-estimators such as Leclerc and Lorentzian when the thresholds are all set to 1.

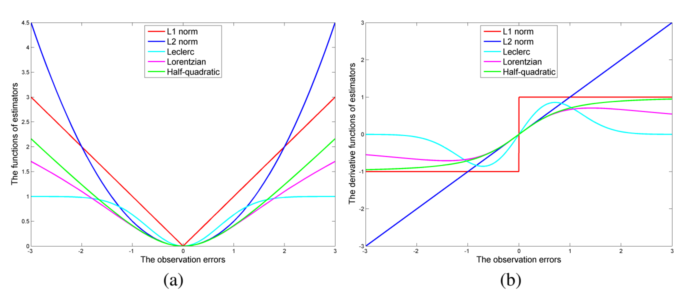


Fig. 1. Error norms. (a) The norm functions of L_1, L_2, Leclerc, Lorentzian and half-quadratic estimation, (b) Their corresponding derivative norm functions.

As shown in Fig. 1, although the Leclerc and Lorentzian could fit L_1 and L_2 norm adaptively according to different inputs, they both have extreme points which make them non-monotonic. Unlike the Leclerc and Lorentzian estimators, the half-quadratic estimation is monotonically increasing. Besides, when the observation error is small, the derivative of half-quadratic function performs like L_2 norm. Subsequently, with the increase of the observation error, the function gradually performs like L_1 norm to suppress outliers. Our adaptive fidelity term is defined as

$$\hat{\mathbf{X}} = arg\min_{\mathbf{X}} \sum_{k=1}^{K} \alpha_k \sqrt{\alpha_k^2 + (\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} - \mathbf{Y}_k)^2}. \quad (6)$$

For each low-resolution frame, α_k is adaptively determined according to the averaged observation error which is defined as $E_k = \|\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X}_0 - \mathbf{Y}_k\|_1/L$, where $\mathbf{X}_0$ is the initial HR estimation and L stands for the total number of pixels in each LR frame. In general, E_k has a small value when the estimation of HR image is accurate. In this case, the observation error fits the Gaussian distribution. The parameter α_k should be large to perform like L_2 norm. In contrast, for those LR frames with outliers and mis-registrations, E_k is large. The parameter α_k should be small to perform like L_1 norm to suppress these kinds of errors. Therefore, we define that α_k is inversely proportional to E_k as

$$\alpha_k = \frac{\max(E_k)}{E_k}. \quad (7)$$

B. Difference Curvature Based BTV Regularization Term

Traditional regularization terms have limited ability to distinguish image edges from noise. Chen *et al.* [11] proposed a new edge indicator called Difference Curvature to distinguish them effectively. It motivates us to combine the traditional BTV regularization with this new edge indicator to suppress noise and preserve edges adaptively. The definition of Difference Curvature is

$$\mathbf{D} = ||\mathbf{I}_{\eta\eta}| - |\mathbf{I}_{\xi\xi}||, \quad (8)$$

$$\mathbf{I}_{\eta\eta} = \frac{\mathbf{I}_x^2\mathbf{I}_{xx} + 2\mathbf{I}_x\mathbf{I}_y\mathbf{I}_{xy} + \mathbf{I}_y^2\mathbf{I}_{yy}}{\mathbf{I}_x^2 + \mathbf{I}_y^2}, \quad (9)$$

$$\mathbf{I}_{\xi\xi} = \frac{\mathbf{I}_y^2\mathbf{I}_{xx} - 2\mathbf{I}_x\mathbf{I}_y\mathbf{I}_{xy} + \mathbf{I}_x^2\mathbf{I}_{yy}}{\mathbf{I}_x^2 + \mathbf{I}_y^2}, \quad (10)$$

Table I shows the performance of $\mathbf{I}_{\eta\eta}$, $\mathbf{I}_{\xi\xi}$ and $\mathbf{D}$ in various areas of a distorted image.

TABLE I
THE PERFORMANCE OF $\mathbf{I}_{\eta\eta}$, $\mathbf{I}_{\xi\xi}$ AND $\mathbf{D}$ IN VARIOUS AREAS

Various areas	$\mathbf{I}_{\eta\eta}$	$\mathbf{I}_{\xi\xi}$	$\mathbf{D}$
Edge	Large	Small	Large
Flat	Small	Small	Small
Isolated noise	Large	Large	Small

In Table I, $\mathbf{I}_{\eta\eta}$, $\mathbf{I}_{\xi\xi}$ and $\mathbf{D}$ are normalized within $[0, 1]$. 'Large' means that the parameter value is larger than 0.5. And 'Small' means that the parameter value is smaller than 0.1. The parameter values between 0.1 and 0.5 are not defined in our algorithm. In general, $\mathbf{I}_{\eta\eta}$ has large value in noise and edge areas but $\mathbf{I}_{\xi\xi}$ only has large value in noise areas. The new edge indicator $\mathbf{D}$ takes the advantage of the difference between them. After subtracting $|\mathbf{I}_{\xi\xi}|$ from $|\mathbf{I}_{\eta\eta}|$, the indicator $\mathbf{D}$ only has large value in edge areas. Therefore, $\mathbf{D}$ has good ability to distinguish edges from noise. After above analysis, our proposed Difference Curvature based BTV regularization (DCBTV) could be formulated as

$$\Upsilon_D(\mathbf{X}) = \sum_{l=-P}^{P} \sum_{m=-P}^{P} \beta^{|m|+|l|}\mathbf{W}_D\left\|\mathbf{X} - \mathbf{S}_x^l\mathbf{S}_y^m\mathbf{X}\right\|_1, \quad (11)$$

where $\mathbf{W}_D$ is the weight matrix and defined as

$$\mathbf{W}_D = \frac{1}{w + \sqrt{\frac{\mathbf{D}}{D_{max}}}}, \quad (12)$$

where w is a positive constant which is set to 0.5 in our experiment and D_{max} is the maximum value of $\mathbf{D}$.

Equ. (13) describes the minimization function of the whole framework.

$$\begin{aligned}\hat{\mathbf{X}} =& arg\min_{\mathbf{X}} \sum_{k=1}^{K} \alpha_k \sqrt{\alpha_k^2 + (\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} - \mathbf{Y}_k)^2} + \\ & \lambda \sum_{l=-P}^{P} \sum_{m=-P}^{P} \beta^{|m|+|l|}\mathbf{W}_D\left\|\mathbf{X} - \mathbf{S}_x^l\mathbf{S}_y^m\mathbf{X}\right\|_1,\end{aligned} \quad (13)$$

where λ is the trade-off parameter to control the balance between the fidelity and regularization term.

In order to solve this minimization function, the Scaled Conjugate Gradients (SCG) is used to find the optimized $\hat{\mathbf{X}}$ and the termination criterion is set to $\eta_t = 10^{-3}$ in our experiment. $f'(\mathbf{X})$ is the first-order derivative function of Equ. (13) with respect to $\mathbf{X}$ which is formulated as

$$\begin{aligned}f'(\mathbf{X}) =& \sum_{k=1}^{K} \frac{\alpha_k(\mathbf{D}\mathbf{B}_k\mathbf{M}_k)^T(\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} - \mathbf{Y}_k)}{\sqrt{\alpha_k^2 + (\mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{X} - \mathbf{Y}_k)^2}} + \lambda \sum_{l=-P}^{P} \\ & \sum_{m=-P}^{P} \beta^{|m|+|l|}\mathbf{W}_D(\mathbf{I} - \mathbf{S}_y^{-m}\mathbf{S}_x^{-l})sign(\mathbf{X} - \mathbf{S}_x^l\mathbf{S}_y^m\mathbf{X}),\end{aligned} \quad (14)$$

where $\mathbf{I}$ is an identity matrix. For convenience, $\mathbf{D}\mathbf{B}_k\mathbf{M}_k$ can be regarded as a system matrix $\mathbf{W}_k$ proposed in [12].

IV. Experimental Results

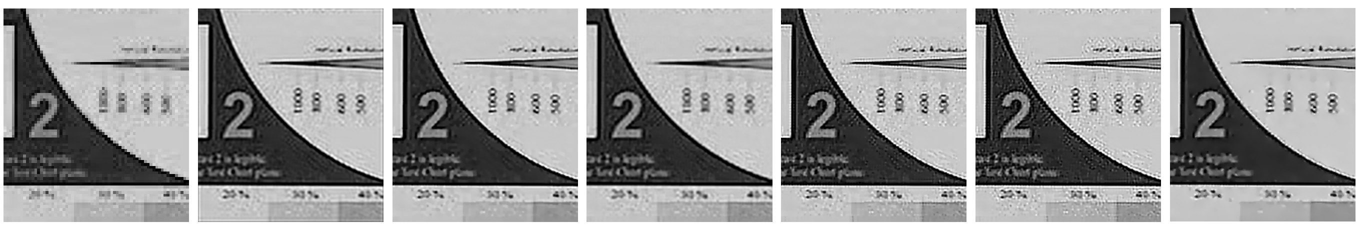
In this section, we use both synthetic and real data to illustrate the performance of our proposed algorithm. Due to space limitation, we only give the results of four sets. The synthetic data was generated by a HR frame and the real data was provided by MDSP dataset [13]. For the synthetic data, the HR image was displaced by random translation matrices and rotation matrices to generate 16 frames. The displaced HR frames were blurred by a 4×4 Gaussian kernel with $\sigma = 0.4$ and then subsampled with factor of $r = 2$. Then we corrupted them with mixed noises containing Gaussian noise ($\sigma_G = 0.02$) and *Salt&Pepper* noise ($\sigma_{SP} = 0.02$). In order

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 2. Super-resolution results for the corrupted *'Cameraman'* image with mixed noise ($r = 2$). (a) Ground truth, (b) LR image (first frame), (c) L_2 + Tikhonov [9] (PSNR:24.77,SSIM:0.58), (d) L_2 + BTV [4] (PSNR:25.30,SSIM:0.76), (e) L_1 + BTV [4] (PSNR:27.38,SSIM:0.84), (f) BEP [8] (PSNR:28.77,SSIM:0.87), (g) IRWSR [5] (PSNR:28.09,SSIM:0.86), (h) Proposed (PSNR:**29.41**,SSIM:**0.88**).

(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3. Super-resolution results for the corrupted *'Lena'* image with mixed noise ($r = 2$). (a) Ground truth, (b) LR image (first frame), (c) L_2 + Tikhonov (PSNR:27.33,SSIM:0.91), (d) L_2 + BTV (PSNR:29.43,SSIM:0.94), (e) L_1 + BTV (PSNR:29.69,SSIM:0.94), (f) BEP (PSNR:30.77,SSIM:0.96), (g) IRWSR (PSNR:31.98,SSIM:0.97), (h) Proposed (PSNR:**33.19**,SSIM:**0.98**).



(a) (b) (c) (d) (e) (f) (g)

Fig. 4. Super-resolution results for *'Adyoron'* data ($r = 3$). (a) LR image (first frame), (b) L_2 + Tikhonov, (c) L_2 + BTV, (d) L_1 + BTV, (e) BEP, (f) IRWSR, (g) Proposed.

(a) (b) (c) (d) (e) (f) (g)

Fig. 5. Super-resolution results for *'Book'* data ($r = 3$). (a) LR image (first frame), (b) L_2 + Tikhonov, (c) L_2 + BTV, (d) L_1 + BTV, (e) BEP, (f) IRWSR, (g) Proposed.

to simulate inaccurate estimation of subpixel movement, the reconstruction procedure did not handle the rotation transformation, which introduces the displacement error on purpose. In the DCBTV regularization, β is set to 0.6 and P is set to 2. The assessment metrics we use to compare our proposed algorithm with others are PSNR (dB) and SSIM. Fig. 2 and Fig. 3 show that our proposed algorithm could effectively suppresse the mixed noises and displacement errors. Meanwhile, it preserves the more texture information than other state-of-the-art algorithms. The PSNR and SSIM values also demonstrate the outperformance of our proposed algorithm. For the real data provided by MDSP dataset, the camera motion and the PSF kernel are unknown. We assume that the real PSF kernel is a 4×4 Gaussian kernel with $\sigma = 0.4$. For the motion estimation, the ECC [14] method is employed to align the LR frames. Super-resolved *Adyoron* and *Book* images are shown in Fig. 4 and Fig. 5 respectively under $r = 3$. Compared with other algorithms, our proposed algorithm has less noise and preserves more detail information in edge areas.

V. Conclusion

In this paper, we proposed a robust multi-frame super-resolution algorithm with adaptive norm choice and regularized by the Difference Curvature based BTV regularization (DCBTV). In our experimental results, both synthetic data and real data are tested to illustrate the performance of our algorithm. Due to the improvements of fidelity term and regularization term, our final results have better quality in visual comparison and higher values in PSNR and SSIM compared with other state-of-the-art methods.

REFERENCES

[1] R. Tsai and T. S. Huang, "Multiframe image restoration and registration," *Advances in Computer Vision and Image Processing*, vol. 1, no. 2, pp. 317–339, 1984.

[2] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space-invariant blur," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1187–1193, 2001.

[3] M.-C. Chiang and T. E. Boult, "Efficient super-resolution via image warping," *Image and Vision Computing*, vol. 18, no. 10, pp. 761–771, 2000.

[4] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[5] T. Köhler, X. Huang, F. Schebesch, A. Aichert, A. Maier, and J. Hornegger, "Robust multiframe super-resolution employing iteratively re-weighted minimization," *IEEE Transactions on Computational Imaging*, vol. 2, no. 1, pp. 42–58, 2016.

[6] V. Patanavijit and S. Jitapunkul, "A robust iterative multiframe super-resolution reconstruction using a huber bayesian approach with huber-tikhonov regularization," in *IEEE Conference on International Symposium on Intelligent Signal Processing and Communications*, 2006, pp. 13–16.

[7] L. Yue, H. Shen, Q. Yuan, and L. Zhang, "A locally adaptive L_1- L_2 norm for multi-frame super-resolution of images with mixed noise and outliers," *Signal Processing*, vol. 105, pp. 156–174, 2014.

[8] X. Zeng and L. Yang, "A robust multiframe super-resolution algorithm based on half-quadratic estimation with modified BTV regularization," *Digital Signal Processing*, vol. 23, no. 1, pp. 98–109, 2013.

[9] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1646–1658, 1997.

[10] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Transactions on Image Processing*, vol. 6, no. 2, pp. 298–311, 1997.

[11] Q. Chen, P. Montesinos, Q. S. Sun, P. A. Heng *et al.*, "Adaptive total variation denoising based on difference curvature," *Image and Vision Computing*, vol. 28, no. 3, pp. 298–306, 2010.

[12] L. C. Pickup, D. P. Capel, S. J. Roberts, and A. Zisserman, "Overcoming registration uncertainty in image super-resolution: maximize or marginalize?" *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, pp. 1–14, 2007.

[13] S. Farsiu, "MDSP super-resolution and demosaicing datasets," https://users.soe.ucsc.edu/~milanfar/software/sr-datasets.html.

[14] G. D. Evangelidis and E. Z. Psarakis, "Parametric image alignment using enhanced correlation coefficient maximization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1858–1865, 2008.