Super-Resolution Image Reconstruction Based on Guidance Image

Imran Khan, Faisal Mufti
Center for Advance Studies in Engineering (CASE)
Islamabad, Pakistan
Pakistan
imi_case@yahoo.com, faisal_mufti@yahoo.com

ABSTRACT: *Super-resolution framework uses multiple noisy low resolution images from camera to generate a higher resolution image that has better spatial resolution than any of the available low resolution images. Super-resolution is an ill-posed problem. A inherent difficulty is the challenge of inverting the image observation model without amplifying the effect of noise in the measured data. Classically, the issue is addressed by incorporating regularization in the cost function to constraint the space of solutions. The main focus of this paper is to develop a regularization function that would preserve edges with improved resolution of super-resolved image. The proposed method is also compared with the classical super-resolution methods and experimental results show the effectiveness and robustness of this method both visually and quantitatively.*

Keywords: Image Reconstruction, Super-resolution, Image models, Image Resolution

Received: 2 September 2014, Revised 7 October 2014, Accepted 13 October 2014

© 2014 DLINE. All Rights Reserved.

1. Introduction

Applications in computer vision domain are dependent on image processing algorithms. These algorithms are often limited in performance due to the characteristic of imaging sensors with respect to the quality of an image. Superresolution (SR) technique is used for reconstruction of highresolution (HR) image from set of low-resolution (LR) images that are typically taken at a different instant in time or from slightly different viewpoints [1]. The images captured by low resolution cameras are often degraded version of the original scene. These degradations may be due to optical blur from the lens PSF and pixel integration, decimation due to the limitation on camera sensor and signal quantization, motion blur due to limited shutter speed and noise that occurs due to signal processing in the camera circuitry. Consequently, the finer details present in the scene are lost during image acquisition process.

Papoulis [2], first discussed SR for recovering igh frequency component from a single image. Later Tsai and Hung [3] exploited aliasing in low resolution images to construct super-resolved image using frequency domain. However, the method fail to address the inherent issue of blurring in low resolution images. Iterative backprojection method is used in [4],

wherein a guess of the HR output image is updated according to the error between the observed and the LR images obtained by simulating the imaging process. But back-projection methods can be used only for those blurring processes for which such an operator can be calculated. Projection onto convex sets (POCS) based methods are described in [5], [6].

Although these methods are quite simple, the interpretation of the result in terms of its spectral content is very difficult. Maximum Likelihood (ML) estimation has also been applied to SR reconstruction [7]. ML estimation is a special case of Maximum a posterior (MAP) estimation (no prior term). As prior information is essential for the solution of ill-posed inverse problems, set theoretic methods, especially the method of projection onto convex sets, are popular as they are simple and allow convenient inclusion of a priori information as regularization term.

Spatial domain methods [4], [8] have also be explored by researchers. A main area of focus in SR is to address the illposed nature of the problem by incorporating a regularization as a priori. This helps to stabilize the effect of noise in resultant image. Several techniques using different priors have been discussed in literature [9]. In [10], edge direction interpolation was presented to preserve edges. Farsiu et al. [11] used bilateral total variation for regularization to address noise and blurring. However, this filter is effective in many situations, it may have unwanted gradient reversal artifacts [12], [13] near edges.

In this paper, we propose a SR framework using a Guidance image [14] regularization for edge enhancement and noise removal. In our approach a Guidance image could be input image itself or any other low resolution image of the same scene. In the proposed framework, the final solution is guided by the regularization term for noise and discontinuities. As a result, the final solution achieves better noise reduction, edge enhancement and improved overall spatial resolution.

The outline of this paper is as follows. The SR observation model is described in Section 2. The proposed SR estimator function is outlined in Section 3. Experimental setup and results on simulated data set is presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Super-resolution Observation Model

Limited bandwidth and discretization of data in digital devices results in low observation of an otherwise infinite resolution scene. The observed LR image obtained from the camera is related to a HR image representing the actual scene through the following observation model, here, the k^{th} *LR* input image is denoted by $\overline{Y}_k$, where $k = 1,2....n$, n is the total number of observed images. Each LR image is arranged in lexicographical form of size $N_1N_2 \times 1$ pixels. The resolution enhancement factor in horizontal and vertical directions is represented by L_1 and L_2 respectively. For the sake of simplicity it is assumed that $L = L_1 = L_2$ without the loss of generality. HR image is denoted by $\overline{X}$ having size $N \times 1$, $N = L^2N_1N_2$. Geometric wrap, representing transformation between frames of images is given by F_k translational matrix of size $N \times N$, B_k is the blurring matrix, representing optical blur of camera having size $N \times N$, D_k is the decimation matrix, representing scale of *LR* image of size $N_1N_2 \times N$. Finally η_k stands for additive zero mean Gaussian noise in the k^{th} measurement. In the absence of any prior knowledge of noise, it is normally assumed white noise and treated as identity matrix [3]. The SR observation model is shown in Figure 1, showing all factors effecting the formation of image.

$$Y_k = D_k B_k F_k \overline{X} + \overline{\eta}_k \text{ for } 1 \le k \le n \tag{1}$$

3. Super-resolution Estimator Function

A fundamental issue in SR reconstruction is the presence of noise in LR images of varying signal-to-noise ratio (SNR). This results in significant edge deterioration of super-resolved image in the absence of any noise removal or filtering technique.

A classical restoration algorithm is based on the assumption that the noise process $\overline{\eta}$ is uncorrelated and has a uniform variance in all observed LR images in (2), the maximum likelikhood estimate of a HR image, $\overline{X}$, can be estimated by minimizing the following cost function,

$$\Phi(X) = argmin \left[\sum_{k=1}^{n} \| D_k B_k F_k \overline{X} - \overline{Y}_k \|^2 \right]. \tag{2}$$

A regularization term compensates the missing measurement information with some general prior information about the desirable solution, and is usually implemented as a penalty factor in the generalized minimization cost function as in (6).

3.1 Edge-Preserving and Enhancing Regularization

Our proposed algorithm incorporates an edge preserving and enhancing regularization term in the from of improved Guidance image [14]. Our algorithm not only constructs a super-resolved image from set of overlapping LR images, but also removes unwanted noise components appeared during reconstruction process while enhancing edges.

We introduce a robust regularizer in the form of Guidance image [14] by considering the fact that it perform well against noise and preserves edges. We consider each LR image as a Guidance image for the currently processed image by utilizing the contents of Guidance image I for noise removal and edge preservation. The smoothness of image and preservation of its contents is achieved by introducing following regularization term denoted by γ in cost function (2) and is defined as

$$\gamma = \frac{1}{|\omega|} \sum_{k:i \in \omega_k} (C_{k1} I_i + C_{k2}), \tag{3}$$

this local linear model ensures that γ has an edge only if I has an edge, we assume that γ is a linear transform of I in a window ω_k centered at the pixel k, where,$|\omega|$ is the number of pixels, P is the input to the filter and C_{k1}, C_{k2} are linear coefficients and are assumed to be constant in window ω_k and are defined as

$$C_{k1} = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i P_i - \mu \overline{P}_k}{\sigma_k^2 + \epsilon} \tag{4}$$

here, $\overline{P}_k = \frac{1}{|\omega|} \sum_{i \in \omega_k} \overline{P}_i$ is the mean of P in ω_k also μ_k and σ^2 are the mean and variance of I in ω_k . ε is the internal pixel adjustment parameter used in reconstruction process and is mainly constant parameter and C_{k2} is defined as

$$C_{k2} = P_k - C_{k1} \mu_k, \tag{5}$$

the final cost function after incorporating the prior term (3) in (2) becomes

$$\Phi(X) = argmin \left[\sum_{k=1}^{n} \| D_k B_k F_k \overline{X} - \overline{Y}_k \|^2 + \lambda \| \gamma \|^2 \right], \tag{6}$$

here the first term on the left hand side is the error norm and the second term on the right hand side is regularization term with λ as regularization parameter in (6) and its values are adjusted accordingly. Let $A_k = D_k B_k F_k$, $\overline{X}$ is HR image to be estimate and $\overline{Y}_k$ are the observed LR images.

$$\Phi(X) = argmin \left[\sum_{k=1}^{n} \| A_k \overline{X}_k - \overline{Y}_k \|^2 + \lambda \| \gamma \|^2 \right]. \tag{7}$$

3.2 Super-Resolution As Optimization Problem

The resultant minimization problem (7) can easily be optimized using standard Conjugate gradient method [15]. Conjugate gradient method converges to the optimum solution X after few iterations, the optimization process is shown below.

Suppose C_0 is an intial vector.

Calculate $C_0 = Y_k - A_k C_0$ and $P_1 = R_0$.

For $k = 1, 2 \ldots$ repeate the following iterations:

- $\alpha_k = \frac{\|R-1\|^2}{P_k . (AP_k)}$
- $C_k = C_{k-1} + \alpha_k P_k$.
- $R_k = Y_k - A_{reg} C_k$.
- $n_k = \frac{\|R_k\|^2}{\|R_k - 1\|^2}$
- $P_{k+1} = R_k + n_k P_k$.

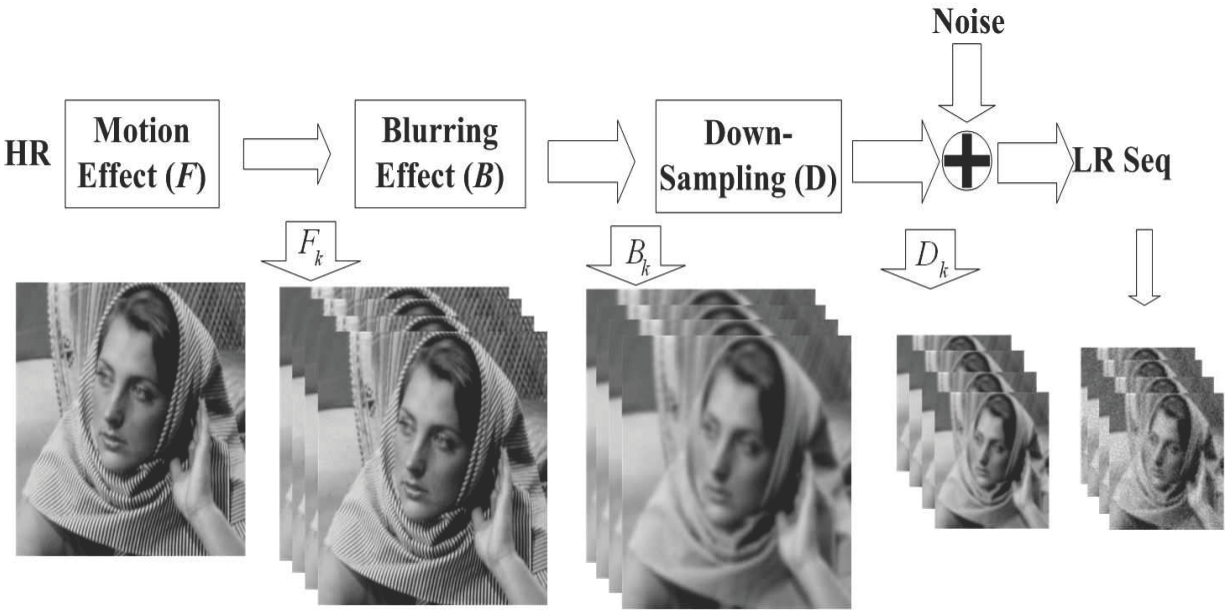


Figure 1. Super-Resolution Observation Model

(a) (b) (c)

(d) (e) (f)

Figure 2. Barbara image used to test the proposed SR framework (a) Ground Truth (b) LR (c) SRR (d) ZMT (e) RSR (f) PROPOSED.

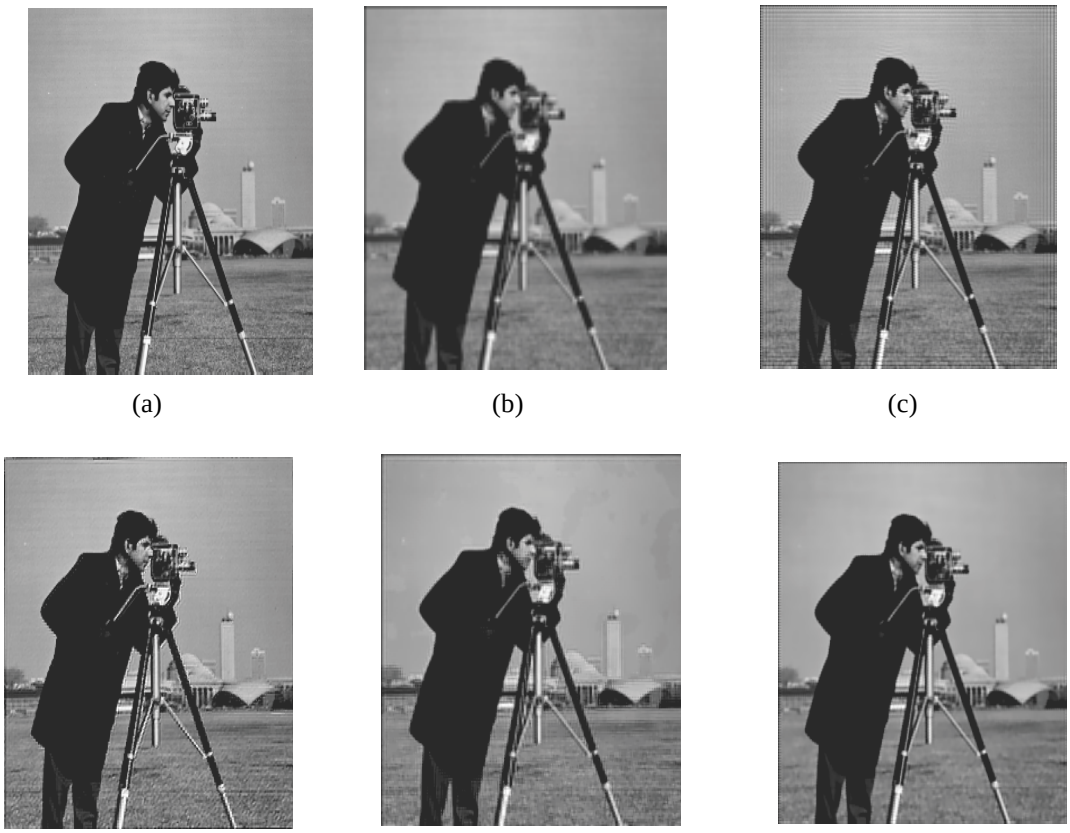
• If $||R_k|| < \varepsilon$, then stop else iterations continue, where is threshold value, n is the step size, R is the residue and P is the direction.

• Finally we get $X = C$ (here X = super-resolved image).

4. Experimental Results

In this section, the performance of the proposed SR framework is evaluated using a set of simulated sequence of images.

All the original images used are of size 256x256 and taken from USC image database [16], each image is used to generate a sequence of 15 synthetically generated LR images. Noise is added by translating LR images in various directions. The ground truth image shown in Figure 2(a) is shifted by one or two pixels on the HR grid in various directions, images undergoes



(a) (b) (c)

(d) (e) (f)

Figure 3. Cameraman image used to test the proposed SR framework (a) Ground Truth (b) LR (c) SRR (d) ZMT (e) RSR (f) PROPOSED.

(a) (b) (c)

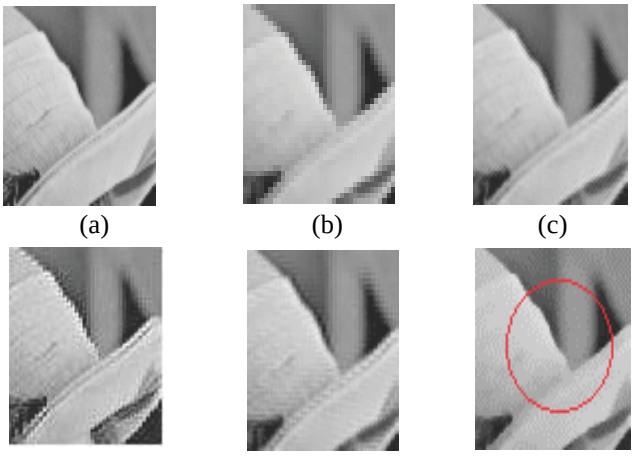
(d) (e) (f)

Figure 4. Lena image used to test the proposed SR framework (a) Ground Truth (b) LR (c) SRR (d) ZMT (e) RSR (f) PROPOSED.

decimations with a factor of 2, the shifted and decimated images were filtered using 3x3 Guassian filter to produce the blurring effect. The LR images were further degraded by additive white Gaussian noise. Images used in this experiment have different characteristics for example Lena image have sharp edges with texture variations whereas the Cameraman image contains smooth regions.

Three methods are used for comparison and results are shown below. These methods include robust SR (ZMT) [17] based on back projection with median filtering, SR reconstruction (SRR) [6], based on a hybrid of ML and POCS using Laplacian prior and the Robust SR (RSR) based on bilateral total variation priors [11]. Qualitative results are shown in Figure 2 to Figure 5, where the proposed method confirms rapid convergence and high Peak signal-to-noise ratio (PSNR). Figure 2(a) is the ground truth "*Barbara*" image, the LR image is shown in Figure 2 (b), Figure 2(c) is the output of approach presented in [6].

Figure 2(d) and (e) present the results of ZMT [17] and RSR [11] respectively. Figure 2 (f) present the obtained SR results using proposed method. Figure 3(c) is the output of approach presented in [6], Figure 3(d) and (e) present the results of ZMT [17] and RSR [11] respectively. Figure 3 (f) presented the output of proposed technique. Figure 4(a) is the ground truth Lena image, Figure 4(c) is the output of approach presented in [6], Figure 4(d) and (e) present the results of ZMT [17] and RSR [11] respectively. Figure 4 (f) presented the output of proposed technique, Figure 5 is the cropped version of Figure 4 to closely analyze the results. Figure 5(f) is the result of proposed method and it indicates good representation of original HR image, both in terms of preserving gray scale values and retaining the image structure. The region inside the red circle shows the edges of various regions remain sharp and intact, with reduce noise compared with other methods shown in Figure 5(c), (d) and (e). The results also reflect effectiveness of the proposed algorithm for removal of unwanted gradient reversal artifacts near edges that appears in other methods [11] during detail enhancement.



(a) (b) (c)

(d) (e) (f)

Figure 5. Detailed regions cropped from Figure 4. (a) Ground Truth Image, (b) LR, (c) SRR, (d) ZMT (e) RSR (f) PROPOSED Method. Inside red encircle part, the edges remains sharp and preserved compared with other methods

We have also tested our proposed algorithm using real data set taken from [18]. Alpaca LR image sequence each of is used

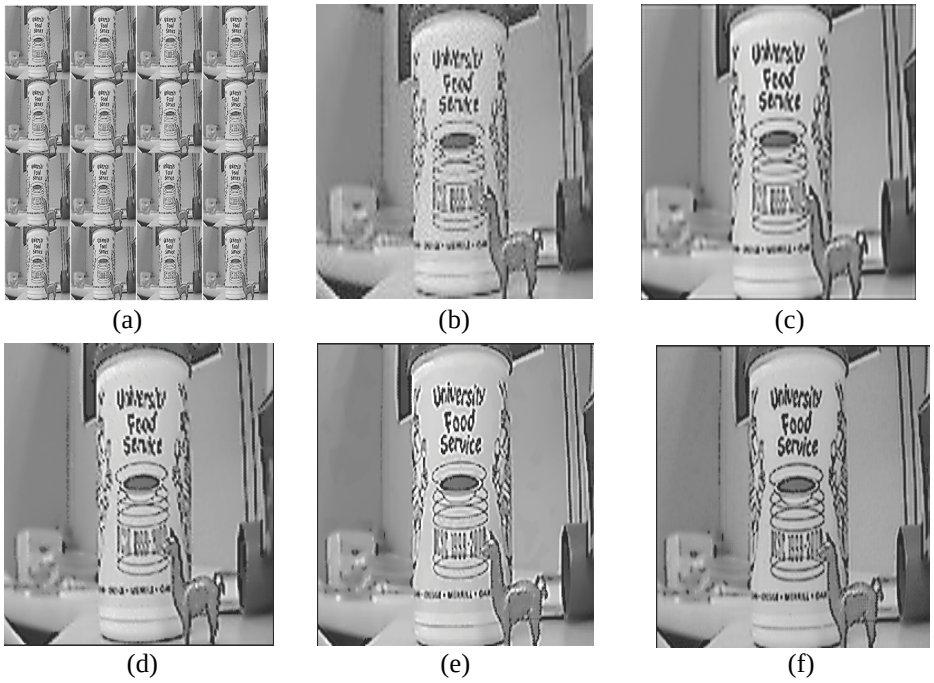


Figure 6. Alpaca image sequence is used to test the proposed SR framework (a) Sequence of LR images (b) Single LR Image (c) SRR (d) ZMT (e) RSR (f) PROPOSED.

to generate SR image. Each frame of this image sequence is of 128X96 dimensions and the resultant SR image is of 256x192. Figure 6(a) is the sequence of LR Alpaca images used in resolution SR reconstruction process. Figure 6(b) is one of the LR image. The result of three algorithms is shown in Figure 6 and the detailed cropped version of these results is shown in Figure 7. The qualitative results clearly illustrate that proposed method well preserve the overall image quality in addition to resolution enhancement compared to other methods. As the ground truth image of Alpaca is not available so it is not possible to calculate PSNR and MSE for this image sequence.

Figure 7. Detailed regions cropped from Figure 6. (a) SRR, (b) ZMT (c) RSR (d) PROPOSED Method

Super-Resolution Methods	**PSNR(db)**
SRR	28.13
ZMT	27.32
RSR	26.47
PROPOSED	29.61

Table 1. PSNR Comparision

The graphical representation of results in terms of PSNR and MSE is shown in Figure 8 and Figure 9 along with iterations. The tabular representation PSNR comparison for Lena image sequence is presented in Table 1 showing PSNR values of proposed method with other methods. Higher values for PSNR corresponds to better image quality with reduce noise component.

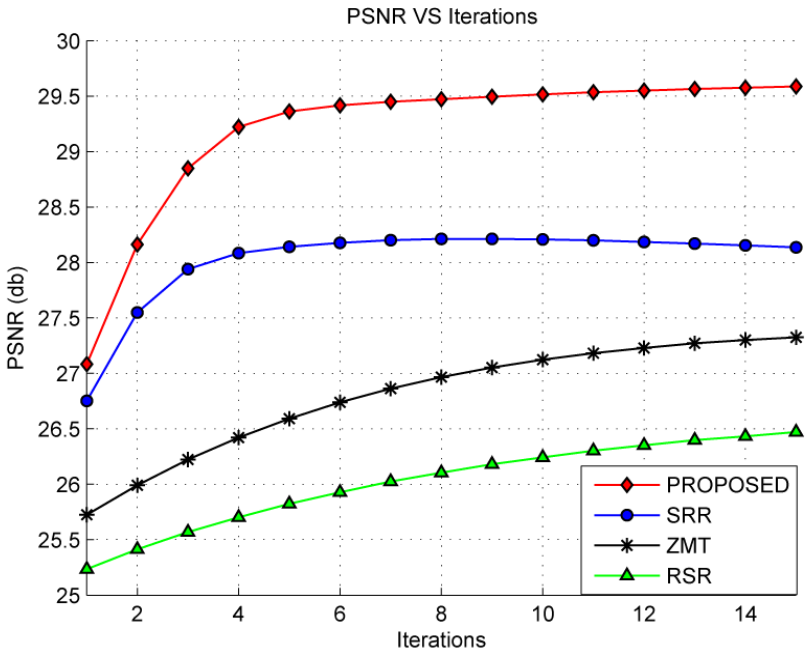


Figure 8. PSNR VS Iterations

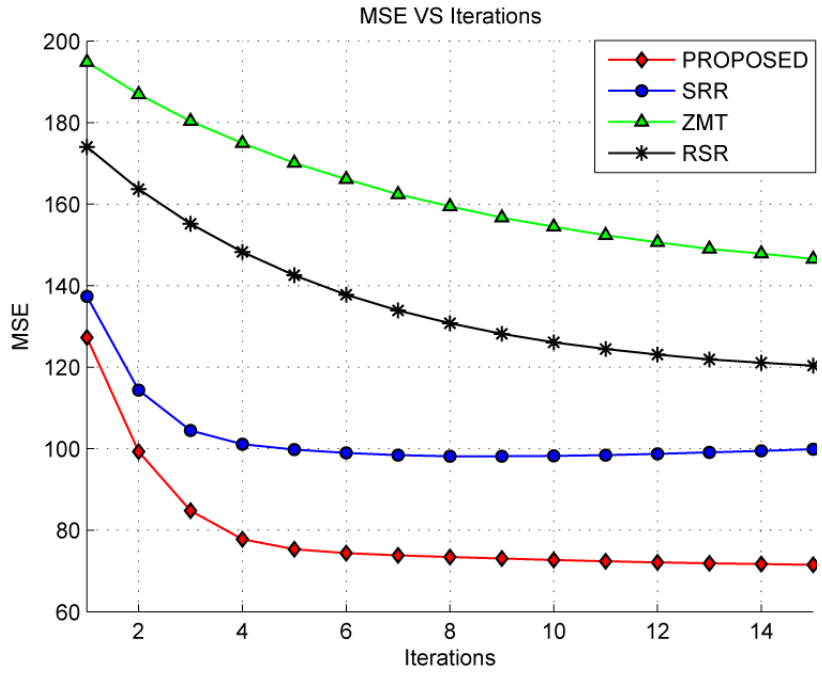


Figure 9. MSE VS ITERATIONS

5. Conclusion

In this paper, we introduced an edge-enhancing SR algorithm using the Guidance image. Our results show that the proposed method preserve edges while smoothing images. The approach used in this paper is robust against noise and also preserve edges during SR image reconstruction. The proposed regularization technique does not perturb the original system, but it

mathematically converts the system to a regularized system using regularization. The results are compared with classical SR techniques and found to be appreciative in terms of PSNR.

References

[1] Segall, C., Molina, R., Katsaggelos, A. (2003). High-resolution images from low-resolution compressed video," Signal Processing Magazine, IEEE, 20 (3), p. 37–48.

[2] Papoulis, A. (1968). Systems and transforms with applications in optics, McGraw-Hill Series in System Science, Malabar: Krieger, 1.

[3] Tsai, R., Huang, T. (1984). Multiframe image restoration and registration, Advances in computer vision and image processing, 1(2), p. 317–339.

[4] Irani, M., Peleg, S. (1990). Super resolution from image sequences, in Pattern Recognition, *In*: Proceedings., 10th International Conference on, 2. IEEE, p. 115–120.

[5] Tekalp, A., Ozkan, M., Sezan, M. (1992). High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration, in Acoustics, Speech, and Signal Processing, ICASSP- 92., IEEE International Conference on, 3. IEEE, p. 169–172.

[6] Elad, M., Feuer, A. (1996). Super-resolution reconstruction of an image, in Electrical and Electronics Engineers in Israel, Nineteenth Convention of. IEEE, p. 391–394.

[7] Tom, B., Katsaggelos, A. (1995). Reconstruction of a high resolution image from multiple degraded mis-registered low resolution images, Ph.D. dissertation, Citeseer.

[8] Chiang, M., Boult, T. (1996). Efficient image warping and super-resolution, in Applications of Computer Vision, WACV'96., *In*: Proceedings 3rd IEEE Workshop on. IEEE, p. 56–61.

[9] Park, S., Park, M., Kang, M. (2003). Super-resolution image reconstruction: a technical overview, Signal Processing Magazine, IEEE, 20 (3), p. 21–36.

[10] Allebach, J., Wong, P. (1996). Edge-directed interpolation, in Image Processing, *In*: Proceedings., International Conference on, 3. IEEE, 1996, p. 707–710.

[11] Farsiu, S., Robinson, M., Elad, M., Milanfar, P. (2004). Fast and robust multiframe super resolution, Image Processing, IEEE Transactions on, 13 (10), p. 1327–1344.

[12] Farbman, Z., Fattal, R., Lischinski, D., Szeliski, R. (2008). Edge-preserving decompositions for multi-scale tone and detail manipulation, in ACM Transactions on Graphics (TOG), 27 (3). ACM, p. 67.

[13] Bae, S., Paris, S., Durand, F. (2006). Two-scale tone management for photographic look, in ACM Transactions on Graphics (TOG), 25(3). ACM, 2006, p. 637–645.

[14] He, K., Sun, J., Tang, X. (2010). Guided image filtering, Computer Vision– ECCV, p. 1–14.

[15] Hestenes, M., Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems.

[16] Weber, A. (1997). The usc-sipi image database version 5, USC-SIPI Report, 315, p. 1–24.

[17] Zomet, A., Rav-Acha, A., Peleg, S. (2001). Robust super-resolution, in Computer Vision and Pattern Recognition, CVPR *In*: Proceedings of the 2001 IEEE Computer Society Conference on, 1. IEEE, 2001, pp. I–645.

[18] Milanfar, P. Mdsp super-resolution and demosaicing datasets, http://www.soe.ucsc.edu/ milanfar/software/sr-datasets.html.