# Convergence Analysis of Real-World Noisy Images in Blind Denoising

Gulsher Ali
**Sukkur IBA University, Pakistan**

Junaid Ahmed
University of Electronic Science and Technology, Chengdu, China

Bilal Tirmizi
Sukkur IBA University, Pakistan

Sharjeel Afridi
Sukkur IBA University, Pakistan

**Abstract**
Most of the image denoising algorithms are tested on standard images with known noise process. It is assumed that the contaminating noise process is Additive White Gaussian Noise (AWGN). Moreover, intensity of the noise is also pre-defined. However, the real noise is much more complex than AWGN. In this paper, the blind image denoising is performed on real-world noisy images from Renoir dataset. The experiment is performed to analyze the convergence (finding optimum solution) performance of sparse or low rank approximation algorithms on removing real camera noise. Performance is evaluated in terms of peak signal-to-noise-ratio (PSNR) and structural similarity index measure (SSIM). Experimental results show that for denoising images with real camera noise the correlation based sparse representation approach keeps finding the most suitable atoms from dictionary and its performance keeps improving with the increasing dictionary subspace when compared to well-known KSVD algorithm.

***Key words:***
*Blind denoising; real-world noisy images; correlation based sparse representation.*

## 1. Introduction

Image denoising is a classic and fundamental problem in the field of computer vision and image processing. Many researchers have proposed numerous algorithms [1]–[4] with approximate optimal results [5]–[7]. Here, the noise is considered to have additive white Gaussian noise (AWGN). However, in the real scenario, this is not the case. In real camera vision, the noise is contributed from different sources (i.e. dark current noise, short noise, thermal noise, etc). Further, the internal camera processing degrades the performance and quality of the images and Gaussian distribution alone is not able to account for it. In some cases, the Gaussian distribution based algorithms are extended for real world scenario, but still their performance is far from better and it remains an open challenge which needs to be addressed for blind denoising of real world images.
With the recent advances in the deep convolutional networks (CNNs) [3], [4]. These algorithms perform well on Gaussian distributed noisy images and fail to perform well on real-world noisy images [8]. As described in [9], the CNNs work on the information from the training data images. The real-world noisy images are quite different from the clean ground-truth ones. Also, the Gaussian noisy images are quite indifferent from the real ones. Due to this reason, the CNN based algorithms fail to perform well in case of real world noisy image.
To tackle this problem, several enhanced models have been proposed. In [10], [11], a correlated Gaussian model is proposed. In [12], [13], a noise model working on the principal of signal and frequency data is proposed. In [14], a model based on the merger of Poisson and Gaussian distribution is proposed for the real world noisy images.
The recent deep neural networks also show promising results. However, early methods have quite limited performance [1], [15], [16]. In [2], authors use a multi-layer perceptron (MLP) approach and achieve good results. In [3], [17], authors utilize the optimization based algorithms for the task of image denoising. In [4], authors utilize the concept of residual learning[18] and batch normalization[19] to achieve good image denoising results. Other CNN based models include RED30 [20], MemNet [21], and FFDNet [22], all of these show promising results. However, their performance on real-world noisy images is quite limited.
The blind denoising of real world images is more difficult than the earlier discussed non-blind models. In [23], authors propose an intensity skellam line estimation method for blind denoising. In [24], authors use the piece wise image smoothing model for the color noise images. In [25], authors model the mixed or unknown noise in the wavelet domain using the weighted $l_1$ *and* $l_2$ sparse regularizer. In [26], authors propose a mixture of Gaussian (MoG) model to estimate the complex noise by using the Bayesian nonparametric approach.
In this paper.
The rest of the paper is as follows; In Section 2, the motivation and problem statement is given. In Section 3, blind denoising framework is defined and algorithm is discussed. In Section 4, the complexity analysis of algorithm is presented. In Section 5, the simulation and results are presented. Section 6 concluded the paper.

Manuscript received February 5, 2019
Manuscript revised February 20, 2019

## 2. Motivation and Problem Statement

Depth of the dictionary subspace needs to be well defined in order to avoid the overfitting and/or underfitting problem. As we know that the image denoising is a convex optimization problem, hence, the dictionary subspace plays important role in denoising performance. This problem becomes more sever for real-world noisy images [27] due to unknown noise process with unknown noise intensity.

Our problem is stated as follows: Given a real-world noisy image corrupted with real camera noise, find its sparse approximation with varying dictionary subspace to analyze the efficiency of low rank image denoising algorithms.

## 3. Blind Denoising Framework of Real-World Noisy Images

### 3.1 Sparse Representation by Correlation Reduction

Let us take a real-world noisy image from Renoir dataset [27]. Note that we consider ground truth image given in [27] as a clean image. Finding the accurate noise level in any noisy image is a challenging task especially if number of samples are not sufficient. Therefore, in order to accurately measure the performance, we vary the dictionary subspace and observe the convergence of the algorithm. Firstly, we divide the clean image $\boldsymbol{x}$ and noisy image $\boldsymbol{y}$ into fully overlapping patches. Then, we find sparse approximation $\boldsymbol{\alpha}$ of each patch using redundant bases called dictionary. After finding sparse representation $\widehat{\boldsymbol{x}} = \boldsymbol{D\alpha}$ , the residual is then given by $\boldsymbol{r} = \boldsymbol{y} - \widehat{\boldsymbol{x}}$. Here $\widehat{\boldsymbol{x}}$ is recovered (denoised) image patch**.** For finding $\boldsymbol{\alpha}$, we employ correlation reduction-based approach [28]. However, in [28], the standard test noisy images with known noise process are denoised. Moreover, dictionary size is also fixed and noisy intensity is predefined. Whereas, in this paper we perform blind denoising (unknown noise process and noise intensity) on real-world noisy images with variable dictionary subspace. So, in this scenario, the sparse approximation method is mathematically formulated as;

$$\widehat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\mathbf{argmin}}\ ||\boldsymbol{\alpha}||_0\ \boldsymbol{s.t} \sum_{\boldsymbol{j}} (|\boldsymbol{u}_{\boldsymbol{j}}^{r} - \boldsymbol{\sigma}^2\boldsymbol{\delta}_{\boldsymbol{j}}|) \leq \boldsymbol{\varepsilon}. \qquad (1)$$

Here $\boldsymbol{u}_j^r$ and $\delta_j$ are the internal autocorrelation of the residual and standard deviation of noise. Internal correlation can mathematically be formulated as;

$$\boldsymbol{U}_{J_1,J_2} = \frac{1}{M}\sum_i \sum_k \boldsymbol{S}_{i,k}\boldsymbol{S}_{i+j,k+j_2} \qquad (2)$$

Internal correlation is determined by selecting a center patch inside residual then finding its correlation with neighboring patches ($j = 0,1,2$ $or$ $3$). Then, all the correlations are summed in absolute sense. Finally, current correlation is removed from previous correlation to find the amount of correlation reduction. The dictionary subspace with highest correlation reduction is selected for sparse approximation. This is to ensure that the autocorrelation of contaminating noise should be similar to the autocorrelation of the residual. It is to note that taking more neighbors increases the complexity, however, it does not improve the denoising performance. Therefore, we kept three neighbors to achieve optimum performance. Note that if $j = 0$ then $\boldsymbol{u}_0^r$ represents the residual power. Hence by minimizing the optimization function (1) one can reduce the residual power as well as non-zero lag correlation to ensure the similarity between residual and contaminating noise.

In (1), we obtain norm-zero i.e., (number of nonzero entries) to set the level of sparsity. After finding sparse representations of all the patches, we update the dictionary as following:

$$\{\widehat{\boldsymbol{\alpha}}, \widehat{\boldsymbol{D}}\} = \underset{\boldsymbol{\alpha},\boldsymbol{D}}{\operatorname{argmin}} \sum_i ||\boldsymbol{\alpha}||_0\ s.t\ \sum_i \sum_j (|\boldsymbol{a}_j^r - \sigma^2\delta_j|) \leq \varepsilon \qquad (3)$$

As (2) is the dual optimization problem, therefore, we initialize the Dictionary with random patches and keep it fix to learn the sparse representation. Once sparse representations of all the patches are achieved, we fix the sparse representation and train the dictionary atoms using (2).

### 3.2 Sparse Representation by Maximum Projection Based Algorithm

K-means Singular Value Decomposition (KSVD) is well known dictionary learning algorithm. Elad et.al [29] performed image denoising via KSVD algorithm. In [29], the standard test noisy images with known noise process are selected for performance evaluation. In [29], the error based Orthogonal Matching Pursuit (OMP) is used to find the sparse approximation. OMP is maximum projection-based algorithm where the atom that produces maximum projection with residual is picked for sparse representation. However, it is shown in [28] that this approach fails at high noise levels or where noise power is greater than power of the clean signal. It is due to fact that, in such scenario, the atom matches the noise process is picked. The sparse representation of KSVD is given by:

$$\widehat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \|\boldsymbol{\alpha}\|_0\ s.t\ \|\boldsymbol{x} - \boldsymbol{D\alpha}\|_2^2 \leq \varepsilon \qquad (4)$$

Here $||.||_2$ and $||.||_0$ represents $\ell_2$ and $\ell_0$ norms, ε represents the approximation error. In optimization problem (4) the denoising is achieved by minimizing the residual power equal to the noise power. This method is effective when noise intensity is known and when noise process is assumed to be AWGN. In this paper, we evaluate the blind denoising performance of [29] on real-world noisy images of Renoir dataset [27] and compare its convergence performance with correlation based denoising approach discussed in previous section.

## 4. Complexity Analysis

In this section, the computational complexity of proposed blind denoising framework is discussed. In this framework we vary the dictionary subspace to find the convergence point and evaluate the stability performance of algorithms. Complexity of algorithm increases with increasing size of the dictionary subspace. Complexity of the algorithm is given by $O(MW_bDLI)$. Where $M$ is patch size, $D$ is number of dictionary atoms, $I$ is the number of iterations and $L$ is number of nonzero entries in each sparse coefficient vector, $W_b$ is the size of dictionary subspace. Then in order to select the most appropriate atom the autocorrelation sequences are calculated. Then these autocorrelation sequences are analyzed. The atom that reduces the autocorrelation most is selected for sparse approximation. It then compares the calculated autocorrelation sequences with that of contaminating noise and determines the atom to be picked. The computational complexity of K-SVD [29] is almost same except that it does not find autocorrelation sequences. Hence, the correlation based approach is slightly complex than KSVD [28].

## 5. Simulation and Results

In this section, we evaluate and compare the performance of correlation-based sparse representation algorithm with maximum projection-based algorithm in blind denoising of real-world images. The performance is evaluated in terms PSNR and SSIM. We randomly select ten real-world noisy images from Renoir data set as shown in Fig. 1. The image is divided into 8×8 fully overlapping patches. Then, we select multiple dictionary subspaces of size 10, 20, 30 and 40 for sparse representation. For each subspace we determine sparse representation and dictionary update stage. We iterate sparse representation and dictionary learning algorithm five times. After completion of five iterations, we reconstruct the image from denoised patches. Finally, we calculate the PSNR and SSIM, by using ground truth images (given in Renoir data set [27]) of all noisy images. We perform the same denoising steps using KSVD [29] algorithm with same parameters for fair comparison.

Fig. 1 Real-World Images from Renoir Dataset [1]

In Fig. 2, we plot the average PSNR results of correlation-based approach and KSVD image denoising algorithm. Fig. 1 clearly indicates that as we increase the size of dictionary subspace the performance of correlation-based algorithm keeps improving for blind denoising of real-world noisy images. Whereas, the performance of KSVD degrades with the increasing dictionary subspace. This is due to fact that residual obtained in correlation-based approach is closer to real camera noise. Whereas, residual obtained from KSVD [29] contains remnants from clean image.

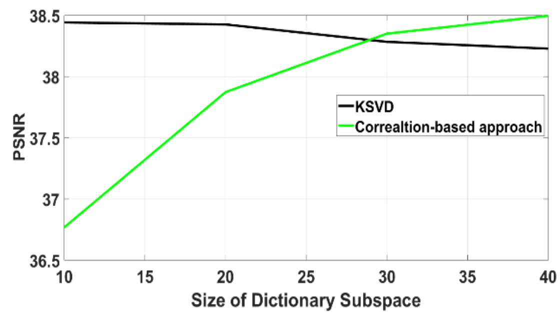


Fig. 2 Average PSNR after blind denoising.

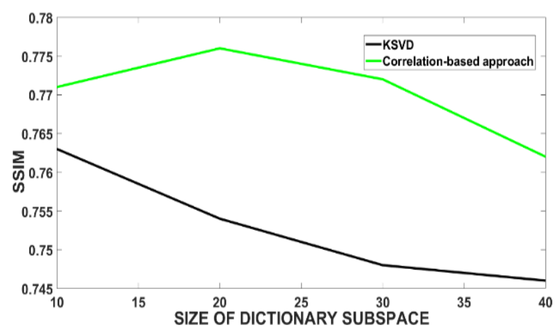


Fig. 3 Average SSIM after blind denoising.

Fig.3 shows the average PSNR results of KSVD and correlation based denoising approach for blind denoising of real-world noisy images from Renoir dataset. It shows that structures are better restored with correlation-based approach. However, as dictionary subspace increases it overfits the data.

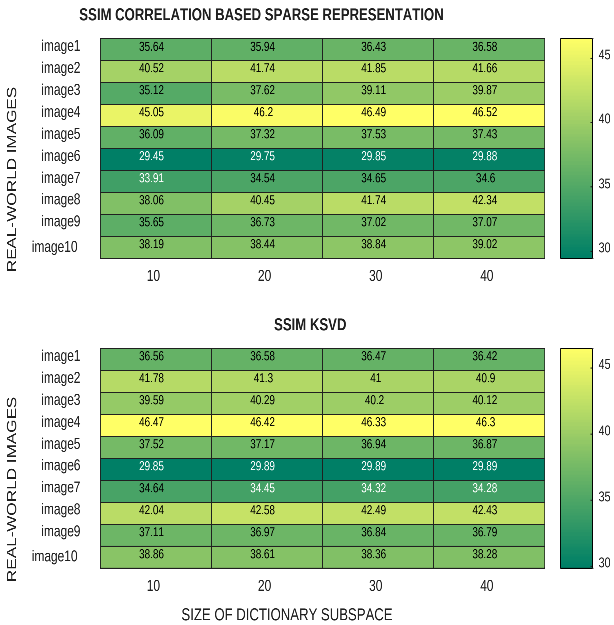


Fig. 4 PSNR of 10 randomly selected Renoir images

Fig. 4 shows the heat map of PSNR values obtained by KSVD [29] and correlation-based approach. The PSNR values of almost all images keeps improving with increasing the dictionary subspace in correlation-based approach. It is due to fact that the correlation-based approach converges slowly and larger the dictionary subspace better the correlation estimation.

## 6. Conclusion

In this paper, blind denoising is performed on real-world noisy images. As real noise is more complex than AWGN, therefore, testing algorithms on AWGN with known intensity does not manifest real performance of the algorithm. Hence, we selected real-world noisy images from well-known data set [27]. We evaluated the convergence performance of low rank approximation methods with varying dictionary subspaces. Correlation-based and maximum projection-based sparse representations are iterated few times with multiple dictionary subspaces. Results show that for real world-noisy images the PSNR performance of correlation-based approach keeps improving with increasing dictionary subspace as compared to maximum projection-based algorithm.

## References

[1] V. Jain and H. Seung, “Natural image denoising with convolutional networks,” Adv. Neural Inf. …, pp. 1–8, 2008.
[2] H. C. Burger, C. J. Schuler, and S. Harmeling, “Image denoising: Can plain neural networks compete with BM3D?,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 2392–2399, 2012.
[3] Y. Chen and T. Pock, “Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1256–1272, 2017.
[4] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising,” IEEE Trans. Image Process., vol. 26, no. 7, pp. 3142–3155, 2017.
[5] P. Chatterjee and P. Milanfar, “Is Denoising Dead ?,” IEEE Trans. Image Process., vol. 19, no. 4, pp. 895–911, 2010.
[6] A. Levin, B. Nadler, F. Durand, and W. T. Freeman, “Patch Complexity , Finite Pixel Correlations and Optimal Denoising,” in Europen Conference on Computer Vision (ECCV), 2012, pp. 73–86.
[7] Y. Romano, M. Elad, and P. Milanfar, “The Little Engine that Could: Regularization by Denoising (RED),” SIAM J. Imaging Sci., vol. 10, no. 4, pp. 1804–1844, 2016.
[8] T. Plötz and S. Roth, “Benchmarking denoising algorithms with real photographs,” Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017–Jan, pp. 2750–2759, 2017.
[9] C. . Martin and M. . Mahoney, “Rethinking Generalization Requires Revisiting Old Ideas: Statistical Mechanics Approaches and Compplex Learning Behavior,” CoRR abs/1710.09553, pp. 1–24, 2017.
[10] J. Portilla, “BLIND NON-WHITE NOISE REMOVAL IN IMAGES USING GAUSSIAN SCALE MIXTURES IN THEWAVELET DOMAIN,” in Proceedings of SPS. Signal Processing Symposium, 2004, pp. 17–20.
[11] J. Portilla, “Full Blind Denoising Through Noise Covariance Estimation Using Gaussian Scale Mixtures in the Wavelet Domain,” in International Conference on Image Processing (ICIP), 2004, pp. 1217–1220.
[12] M. Lebrun, M. Colom, and J. M. Morel, “Multiscale Image Blind Denoising,” IEEE Trans. Image Process., vol. 24, no. 10, pp. 3149–3161, 2015.
[13] M. Lebrun, M. Colom, and J.-M. Morel, “The Noise Clinic: a Blind Image Denoising Algorithm,” Image Process. Line, pp. 1–54, 2015.
[14] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, “Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data,” IEEE Trans. Image Process., vol. 17, no. 10, pp. 1737–1754, 2008.
[15] T. Rabie, “Robust estimation approach for blind denoising,” IEEE Trans. Image Process., vol. 14, no. 11, pp. 1755–1765, 2005.
[16] J. Xie, L. Xu, and E. Chen, “Image Denoising and Inpainting with Deep Neural Networks,” NIPS, pp. 1–9, 2012.
[17] U. Schmidt and S. Roth, “Shrinkage fields for effective image restoration,” Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 2774–2781, 2014.
[18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition Kaiming,” in IEEE International Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[19] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” arXiv:1502.03167v3, 2015.
[20] X. Mao, C. Shen, and Y. Yang, “Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections,” arXiv:1603.09056v2, 2016.
[21] Y. Tai, J. Yang, X. Liu, and C. Xu, “MemNet : A Persistent Memory Network for Image Restoration,” in IEEE International Conference on Computer Vision, 2017, pp. 4549–4557.

[22] K. Zhang, W. Zuo, S. Member, and L. Zhang, “FFDNet : Toward a Fast and Flexible Solution for CNN-Based Image Denoising,” IEEE Trans. Image Process., vol. 27, no. 9, pp. 4608–4622, 2018.
[23] Y. Hwang and J. Kim, “Difference-Based Image Noise Modeling Using Skellam Distribution,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 7, pp. 1329–1341, 2012.
[24] C. Liu et al., “Automatic Estimation and Removal of Noise from a Single Image,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 2, pp. 299–314, 2008.
[25] Z. Gong, Z. Shen, and K.-C. Toh, “Image Restoration with Mixed or Unknown Noises,” Multiscale Model. Simul., vol. 12, pp. 458–487, 2014.
[26] F. Zhu, G. Chen, and P. A. Heng, “From Noise Modeling to Blind Image Denoising Shenzhen Institutes of Advanced Technology , Chinese Academy of Sciences,” 2016 IEEE Conf. Comput. Vis. Pattern Recognit., pp. 420–429, 2016.
[27] J. Anaya and A. Barbu, “RENOIR – A dataset for real low-light image noise reduction,” J. Vis. Commun. Image Represent., vol. 51, no. January, pp. 144–154, 2018.
[28] G. Baloch and H. Ozkaramanli, “Image denoising via correlation-based sparse representation,” Signal, Image Video Process., vol. 11, no. 8, pp. 1501–1508, 2017.
[29] M. Elad and M. Aharon, “Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries Introduction - Issue Revealing,” IEEE Trans. Image Process., vol. 15, no. 12, pp. 3736–3745, 2011.