

IMAGE NOISE ESTIMATION AND REMOVAL CONSIDERING THE BAYER PATTERN OF NOISE VARIANCE

Huanjing Yue¹, Jianjun Liu¹, Jingyu Yang^{1*}, Truong Nguyen² and Chunping Hou¹

¹Tianjin University, ²University of California, San Diego

ABSTRACT

Traditional image denoising methods are designed for Gaussian or Poisson noise, which are not suitable for realistic noise introduced in the complicated imaging pipeline. We observe that, due to the demosaicing process in imaging, the noise variance maps of captured JPEG images are characterized by Bayer patterns. In this paper, we propose a novel noise estimation and removal method based on the Bayer pattern of noise variance maps. There are two key contributions in the proposed method. First, to the best of our knowledge, we are the first to consider the Bayer patterns of noise variance maps in noise estimation and denoising. Second, we extend the state-of-the-art denoising method CBM3D to deal with realistic noise by integrating the estimated noise variance map and Bayer-pattern down-sampling into the denoising process. Experimental results show that the proposed method achieves the best noise estimation performance compared with two state-of-the-art methods. In addition, the denoising performance of CBM3D for realistic noise is significantly improved using the proposed approach and outperforms state-of-the-art blind denoising methods.

Index Terms— Image noise estimation, image denoising, realistic noise, blind denoising

1. INTRODUCTION

Noise is a fundamental problem in many image processing and computer vision tasks. Images captured by digital cameras under high ISO and low light conditions are usually polluted by noise. However this kind of noise cannot be analyzed well using conventional models for Gaussian noise or Poisson noise.

The development of noise models can be categorized as point, (line) curve and map models. Traditional noise estimation methods are designed for the homogeneous white Gaussian noise (WGN), which can be represented by the noise variance (the point model)[1, 2]. Yet, there are many noise sources in real imaging process, such as dark-current noise,

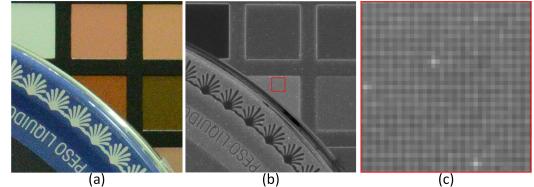


Fig. 1: From left to right: (a) the input noisy image, (b) the noise variance map (in the red channel) of (a), and (c) the selected region of (b).

shot noise, fixed-pattern noise etc, which cannot be modeled by WGN. Consequently, Poisson-Gaussian model is used to model signal dependent noise [3, 4]. The noise level depends on the image intensity. However, this model is only suitable for raw images or linear imaging systems, and is unsuitable for JPEG images, which are the most popular image format used by most consumers and vision researchers. To cope with the non-linear imaging pipeline, the works in [5, 6] utilize a noise level function (NLF) to model the relationship between image intensity and its corresponding noise level. The NLF depends on both the nonlinear camera response function and noise strength (the curve model). Recently, Nam *et al.* proposed a more accurate noise model in terms of noise variance maps, and they claimed that the noise level depends not only on the pixel color but also on the scene surrounding it [7]. Inspired by [7], we also model the noise variance for captured JPEG images by noise variance maps.

For image denoising, many methods have been proposed in the past decades and achieved excellent performance [8, 9, 10, 11]. Nevertheless, most of these methods are designed for WGN removal. Recently, some blind denoising methods are proposed to remove the realistic noise captured by real camera systems [12, 13, 14]. The works in [12, 13] model the noise as signal, scale and frequency dependent and extend the nonlocal-Bayes method to deal with this kind of noise. Zhu *et al.* proposed a low-rank mixture of Gaussian filter to remove noise. These denoising methods are built on traditional Gaussian noise removing models to make them suitable for realistic noise. However, it's still not clear how the denoising methods could fully take advantage of the estimated noise maps.

Fig. 1 (a) presents a typical noisy image captured by Nikon D800 with ISO set to 6400. Fig. 1 (b) is the corre-

This work was supported in part by the National Natural Science Foundation of China under Grant 61672378, Grant 61372084, Grant 61520106002, and in part by the China Scholarship Council. *Corresponding author: yjy@tju.edu.cn.

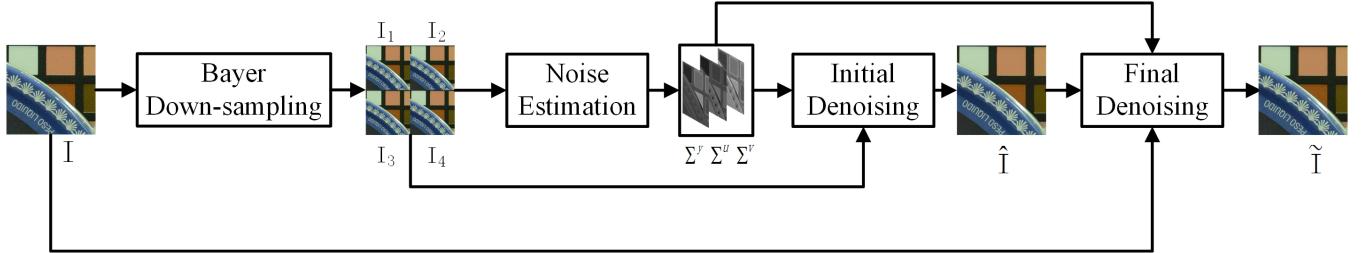


Fig. 2: Framework of the proposed noise estimation and denoising approach. I is the input noisy image and $\{I_1, I_2, I_3, I_4\}$ are the Bayer down-sampled images of I . Σ^y , Σ^u and Σ^v are the noise variance maps for the Y , U , and V channels, respectively. \hat{I} and \tilde{I} are the initial and final denoising results of I .

sponding noise variance map in the red channel for Fig. 1 (a). The ground truth noise variance map is obtained by shooting the same scene for 500 times. One patch, indicated by the red box, of the noise variance map is highlighted in Fig. 1 (c). It can be observed that the noise variance map shows the Bayer-like pattern. The reason is that the demosaicing interpolation process has reduced the noise level for interpolated pixels. Based on this observation, we propose a novel image noise estimation and removal method by taking advantage of the Bayer pattern of the noise variance map.

The contribution of this paper is two-fold. First, to the best of our knowledge, we are the first to propose the Bayer pattern based noise estimation and denoising method for JPEG images. Second, we extend the most popular Gaussian noise removal method CBM3D [15, 8] to deal with realistic noise. In the first stage of CBM3D, we divide the noisy image into four sub-images based on the Bayer patterns and remove their noise separately. Since the reconstructed image by the denoised four sub-images do not satisfy the smoothness of neighboring pixels, we use the whole image to denoise in the second stage.

2. IMAGE NOISE ESTIMATION AND REMOVAL

Fig. 2 depicts the framework of the proposed noise estimation and denoising approach. Given a noisy image I , we first down-sample it into four sub images $\{I_1, I_2, I_3, I_4\}$ according to the Bayer pattern. Then, we propose a non-local similarity based noise estimation method in YUV channel and produce the noise variance maps for the four sub-images separately. Hereafter, we integrate the estimated noise variance map into CBM3D denoising process. In the initial denoising stage of CBM3D, the four sub-images are denoised separately according to the estimated noise variance maps. Then, the four sub-images are recombined into one image \hat{I} as the initial denoising result. In the final denoising stage, the noisy image is denoised with the guidance of \hat{I} and estimated noise variance maps Σ^y , Σ^u and Σ^v .

In the following subsections, the main modules, including noise estimation, initial denoising and final denoising will be discussed in details.

2.1. Noise estimation

The most popular noise variance estimation method in the literature is patch based [4, 6]. However, as shown in Fig. 1 (c), the noise variances in one patch are not uniform, but characterized by Bayer pattern. This inspires us to down-sample the input noisy image into four sub-images $\{I_1, I_2, I_3, I_4\}$, and estimate their noise variance maps separately. Since the estimation process is the same for the four sub-images, we will use I_i to denote the sub-image in the following discussion. We first remove the noise in I_i using CBM3D with a coarse noise variance $\bar{\sigma}^2$. $\bar{\sigma}$ is set according to the ISO values of cameras, namely $\bar{\sigma} = 3\log_2 \frac{\text{ISO}}{100}$. $\bar{\sigma}$ is fixed for the whole image and usually much larger than the true noise level, in order to remove the noise thoroughly. We denote the denoised version of I_i as \bar{I}_i . Compared with I , the noise in \bar{I}_i is greatly reduced. Therefore, we could utilize their difference to estimate the noise variance. We split \bar{I}_i into $k \times k$ blocks with step s . For each block \bar{B} in \bar{I}_i , we first search for its most similar blocks in a local region around \bar{B} . The similar blocks are obtained by minimizing the ℓ_2 distance in the Y channel. We would like to point out that we perform noise estimation and denoising in YUV channels instead of RGB channels, since the YUV channels have less inter-channel correlation than RGB channels. Suppose there are n similar blocks for each query block, then the noise variance of block B in channel c can be calculated as

$$\sigma_B^{c,2} = \text{var}\{B_1^c - \bar{B}_1^c, B_2^c - \bar{B}_2^c, \dots, B_n^c - \bar{B}_n^c\}, \quad c = y, u, v, \quad (1)$$

where B_j is the corresponding noisy version of \bar{B}_j . Here, we obtain the noise variance for each block and merge them together by averaging, generating the final noise variance map Σ_i^y , Σ_i^u , and Σ_i^v for image I_i . Note that, although \bar{I}_i is over-smooth compared with the clean version of I_i , in the experiment we find that it is a good candidate for the noise free version of I_i . The final noise variance map Σ_i^c for image I_i^c is obtained by up-sampling Σ_i^c , $i = \{1, 2, 3, 4\}$ to its original position according to the Bayer pattern.

2.2. Initial denoising

In this paper, we integrate the obtained noise variance maps for the Y , U and V channels into the most popular Gaussian

noise removal method CBM3D to process the noisy image with spatial varying noise variances. A straight forward way to extend CBM3D to handle this kind of noise is adjusting the filtering parameter according to the noise variance in each block, as proposed in [6]. However, as indicated in Fig. 1, the noise variances in a block are not uniform. It's unreasonable to use one fixed parameter to remove the noise in one block. Therefore, similar to the noise estimation process, we down-sample the noisy image into four sub-images and denoise them separately. For each block B in image I_i , we search for its similar blocks in a local region around B , and these similar blocks build a 3D cube B_{3D} . The similar blocks are obtained by minimizing the ℓ_2 norm distance in the Y channel. The denoising blocks are obtained by hard-thresholding of transformed coefficients, denoted as

$$\hat{B}_{3D}^c = \mathcal{T}_{3D}^{-1}(\gamma(\mathcal{T}_{3D}(B_{3D}^c))) \quad c = y, u, v, \quad (2)$$

where T_{3D} is a 3D transform, including a 2D wavelet transform and a 1D Hadamard transform along the third dimension. γ is a hard thresholding operator with the threshold $\delta\lambda^c\sigma_B^c$, where σ_B^c is the mean standard deviation of noise in block B^c . After the 3D inverse transform \mathcal{T}_{3D}^{-1} , we obtain the initial denoising result \hat{B}_{3D}^c .

After obtaining the denoising result for each noisy block in I_i , they are averaged together to produce an initial denoising sub-image \hat{I}_i . Then the four sub images $\{\hat{I}_1, \hat{I}_2, \hat{I}_3, \hat{I}_4\}$ are up-sampled to the original position in image I , generating an initial denoising image \hat{I} .

2.3. Final denoising

Although the noise in the initial denoising result \hat{I} is greatly attenuated, the smoothness of neighboring pixels is destroyed due to the proposed sub-image based denoising scheme. Therefore, in this stage, we propose to denoise image I as a whole with the guidance of \hat{I} . For each block \hat{P} in image \hat{I} ¹, we retrieve its similar blocks and build a 3D cube \hat{P}_{3D} . Correspondingly, we could build another 3D cube P_{3D} by extracting blocks from image I with the same positions as those in \hat{P}_{3D} . Then the final denoising block is obtained by

$$\tilde{P}_{3D}^c = \mathcal{T}_{3D}^{wie^{-1}}(W^c \odot (\mathcal{T}_{3D}^{wie}(P_{3D}^c))) \quad c = y, u, v, \quad (3)$$

where \mathcal{T}_{3D}^{wie} is a 3D transform consisted of a 2D DCT transform and a 1D Hadamard transform along the third dimension. \odot denotes the element-wise multiplication. W^c is the Wiener shrinkage parameter of the c channel, calculated as

$$W^c = \frac{|\mathcal{T}_{3D}^{wie}(\hat{P}_{3D}^c)|^2}{|\mathcal{T}_{3D}^{wie}(\hat{P}_{3D}^c)|^2 + (\beta^c\sigma_P^c)^2} \quad c = y, u, v, \quad (4)$$

where σ_P^c is the mean standard deviation of noise in block P^c . Different from the original Wiener shrinkage parameter

¹To avoid ambiguity, we utilize B to denote the block in sub-images, and utilize P to denote the block in the whole image in this paper.

proposed in CBM3D, we introduce a parameter β^c to adjust the weighting parameter. The reason is that the noise in block P^c is not white, and we observe that the noise variance in the low-frequency band is much larger than that in the high-frequency band. Therefore, we introduce β^c , which is larger than 1, to remove the noise in low-frequency band thoroughly.

After obtaining all the denoised blocks \tilde{P} , we integrate them together by averaging, generating the final denoised image \tilde{I} .

3. EXPERIMENTAL RESULTS

The parameters used in this paper are set as follows. The matrix A used to convert from RGB to YUV channels is

$$A = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix}. \quad (5)$$

The block size k and step s in noise estimation is set to 4 and 1, respectively. The parameter λ^c in initial denoising stage is set to 1.3, 2.7, and 3 for Y, U , and V channels. δ is set to 2.7. The parameter β^c is set to 2.5, 6.4, and 7.4 for Y, U , and V channels, respectively. These parameters are set experimentally and fixed for all test images.

In the following subsections, we will demonstrate the effectiveness of proposed method by comparing with state-of-the-art noise estimation and removal methods. We utilize the dataset² shared by the authors of [7]. The detailed illustration of this dataset and our test images are presented in the supplemental material.

3.1. Noise estimation results

We utilize mean square error (MSE) value to measure the distance between the estimated and ground truth noise deviation maps. Let Σ_{gt}^c denote the ground truth noise variance maps in the c channel. The distance D^c is calculated as

$$D^c = \frac{\|\sqrt{\Sigma^c} - \sqrt{\Sigma_{gt}^c}\|_F^2}{N}, \quad (6)$$

where N is the number of pixels in the noise variance map Σ^c , and $\|\cdot\|_F$ represents the Frobenius norm. It's difficult to directly compare our method with previous noise estimation works, since the noise variance values are estimated in different forms. In this paper, we compare with [7] and [6], which are the most related works to ours. For [7], we transform their estimated noise covariance maps in RGB channels to variance maps in YUV channels, as

$$\begin{aligned} \sigma^c &= \sqrt{\phi_A + \phi_B} \\ \phi_A &= a_{c1}^2\sigma_r^2 + a_{c2}^2\sigma_g^2 + a_{c3}^2\sigma_b^2, \\ \phi_B &= 2a_{c1}a_{c2}\sigma_{rg} + 2a_{c1}a_{c3}\sigma_{rb} + 2a_{c2}a_{c3}\sigma_{gb}, \end{aligned} \quad (7)$$

²<http://snam.ml/research/ccnoise>

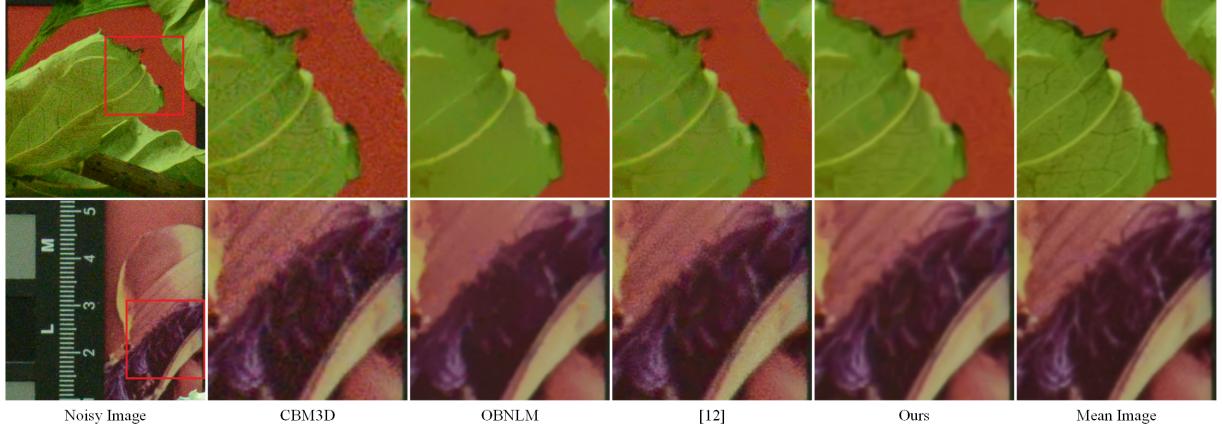


Fig. 3: Comparison of denoising results.

where σ^c represents the standard deviation in the c channel of one pixel and a_{ci} is the transform coefficient $A(c, i)$ defined in Eq. 5. The noise variance estimated in [6] is a NLF depending on the intensity. We map the noise level function to a noise variance map according to the denoised intensity in the Y channel. Since the U and V channels cannot be represented by the NLF, we only compare with [6] in terms of D^y . The results of [7] and [6] are generated by the authors' code.

Table 1 presents the MSE values in Y , U , and V channels for the three methods. We observe that our method consistently outperforms both [7] and [6] for all the three channels.

Table 1: Comparison of the noise estimation results. The best results are highlighted in bold.

Img.	[6]		[7]			Ours		
	D^y	D^{uv}	D^y	D^u	D^v	D^y	D^u	D^v
	PSNR	SSIM	PSNR	SSIM	PSNR	PSNR	SSIM	PSNR
1	2.191	-	7.166	0.533	1.508	0.628	0.109	0.308
2	2.745	-	14.91	0.512	0.995	1.314	0.133	0.299
3	2.791	-	4.549	1.790	0.596	2.511	0.162	0.522
4	2.051	-	2.861	1.424	1.084	1.525	0.218	0.500
5	1.462	-	1.344	1.783	2.114	0.537	0.449	0.539
6	0.840	-	1.508	1.044	1.767	0.584	0.310	0.839
7	1.736	-	5.164	1.327	9.681	0.808	0.948	1.875
8	2.028	-	9.966	0.549	9.292	0.795	0.682	2.325
9	0.603	-	0.933	0.978	0.724	0.397	0.236	1.191
10	0.466	-	0.716	2.271	1.125	0.429	0.201	0.794
Ave.	1.691	-	4.912	1.221	2.889	0.953	0.345	0.919

3.2. Image denoising results

We compare our denoising method with the original CBM3D with the mean ground truth noise level, the Bayesian non-local means (BNLM) algorithm with the ground truth noise covariance maps [7][16], and the blind denoising method [12]. Since the similarity measurement in BNLM is optimized to process realistic noise, we denote its results as OBNLM. Table 2 presents the objective comparison results for the four methods in terms of PSNR and SSIM [17] values. The SSIM results are calculated by averaging the SSIM values in R, G and B channels. It can be observed that the results

of original CBM3D are the worst since it utilizes a fixed noise variance. Compared with OBNLM, although it utilizes the ground truth noise variance maps as input, our method still achieves more than 1 dB gain. Fig. 3 presents the visual comparison results for Image 7 and 8. It can be observed that our method recovers more details and the recovered images are much cleaner compared with the other three methods. More visual results are provided in the supplemental material.

Table 2: Comparison of denoising results in terms of PSNR and SSIM values. The best results are highlighted in bold.

Img.	CBM3D		OBNLM		[12]		Ours	
	σ from GT	σ from GT	σ from GT	σ from EST	-	σ from EST		
		PSNR	SSIM		PSNR	SSIM	PSNR	SSIM
1	38.94	0.980	39.53	0.974	38.78	0.984	40.69	0.988
2	35.91	0.974	35.24	0.951	35.30	0.974	36.94	0.978
3	37.92	0.964	38.71	0.968	38.28	0.976	39.71	0.981
4	36.83	0.963	36.88	0.947	36.74	0.963	38.12	0.973
5	34.46	0.917	36.91	0.946	35.78	0.947	38.00	0.969
6	36.05	0.913	39.17	0.963	37.61	0.952	40.79	0.977
7	32.32	0.861	35.16	0.950	33.07	0.897	35.94	0.957
8	31.39	0.813	34.27	0.902	33.18	0.884	34.84	0.927
9	35.03	0.939	37.74	0.959	37.69	0.962	38.61	0.973
10	35.55	0.957	36.10	0.963	36.72	0.975	36.95	0.981
Ave.	35.44	0.928	36.97	0.952	36.32	0.951	38.06	0.970

4. CONCLUSION

In this paper, we proposed a novel noise estimation and denoising algorithm by considering the Bayer pattern of noise variance maps. Since the noise variances in one block are non-uniform, we divide the input image into four sub-images according to the Bayer pattern. Then we estimate the noise variance maps for the four sub-images separately. In the denoising stage, we extend CBM3D to process realistic noise by integrating the estimated noise variance maps and Bayer down sampling process. Experimental results demonstrate that our method achieves the best noise estimation and denoising results compared with state-of-the-art methods.

5. REFERENCES

- [1] Antonio De Stefano, PR White, and WB Collis, “Training methods for image noise level estimation on wavelet components,” *EURASIP Journal on Applied Signal Processing*, vol. 2004, pp. 2400–2407, 2004.
- [2] Xinhao Liu, Masayuki Tanaka, and Masatoshi Okutomi, “Noise level estimation using weak textured patches of a single noisy image,” in *2012 19th IEEE International Conference on Image Processing*. IEEE, 2012, pp. 665–668.
- [3] Alessandro Foi, Mejdi Trimeche, Vladimir Katkovnik, and Karen Egiazarian, “Practical poissonian-gaussian noise modeling and fitting for single-image raw-data,” *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1737–1754, 2008.
- [4] Xinhao Liu, Masayuki Tanaka, and Masatoshi Okutomi, “Practical signal-dependent noise parameter estimation from a single noisy image,” *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4361–4371, 2014.
- [5] Ce Liu, William T Freeman, Richard Szeliski, and Sing Bing Kang, “Noise estimation from a single image,” in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*. IEEE, 2006, vol. 1, pp. 901–908.
- [6] Jingyu Yang, Ziqiao Gan, Zhaoyang Wu, and Chunping Hou, “Estimation of signal-dependent noise level function in transform domain via a sparse recovery model,” *IEEE Transactions on Image Processing*, vol. 24, no. 5, pp. 1561–1572, 2015.
- [7] Seonghyeon Nam, Youngbae Hwang, Yasuyuki Matsushita, and Seon Joo Kim, “A holistic approach to cross-channel image noise modeling and its application to image denoising,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 1683–1691.
- [8] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *IEEE Transactions on image processing*, vol. 16, no. 8, pp. 2080–2095, 2007.
- [9] Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li, “Nonlocally centralized sparse representation for image restoration,” *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1620–1630, 2013.
- [10] Daniel Zoran and Yair Weiss, “From learning models of natural image patches to whole image restoration,” in *2011 International Conference on Computer Vision*. IEEE, 2011, pp. 479–486.
- [11] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng, “Weighted nuclear norm minimization with application to image denoising,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2862–2869.
- [12] Marc Lebrun, Miguel Colom, and Jean-Michel Morel, “Multiscale image blind denoising,” *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3149–3161, 2015.
- [13] Marc Lebrun, Miguel Colom, and Jean-Michel Morel, “The noise clinic: A blind image denoising algorithm,” *Image Processing On Line*, vol. 5, pp. 1–54, 2015.
- [14] Fengyuan Zhu, Guangyong Chen, and Pheng-Ann Heng, “From noise modeling to blind image denoising,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 420–429.
- [15] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, “Color image denoising via sparse 3d collaborative filtering with grouping constraint in luminance-chrominance space,” in *2007 IEEE International Conference on Image Processing*. IEEE, 2007, vol. 1, pp. I–313.
- [16] Charles Kervrann, Jérôme Boulanger, and Pierrick Coupé, “Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal,” in *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer, 2007, pp. 520–532.
- [17] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.