

000
001
002054
055
056003 **External Patch Group Prior Guided Internal Prior Learning for Real Image
004 Denoising**057
058
059005
006
007060
061
062008
009
010063
064
065011
012
013066
067
068

014

069

Abstract

For image denoising problem, the external and internal priors are playing key roles in many different methods. External priors learn from external images to restore noisy images while internal ones exploit priors of given images for denoising. The external priors are more generative and efficient on recovering structures existing in most images while the internal priors are more adaptive on recovering details existed in given noisy images. In this paper, we propose to employ the external patch group prior of images to guide the clustering of internal patch groups, and develop an external dictionary guided internal orthogonal dictionary learning algorithm for real image denoising. The internal orthogonal dictionary learning process has closed-form solutions and hence very efficient for online denoising. The experiments on standard datasets demonstrate that, that the proposed method achieves better performance than other state-of-the-art methods on real image denoising.

1. Introduction

Most vision systems, such as medical imaging and surveillance, need accurate feature extraction from high-quality images. The camera sensors and outdoor low light conditions will unavoidably bring noise to the captured images. The impact is that the image details will be lost or hardly visible. As a result, image denoising is an essential procedure for the reliability of these vision systems. In the research area, image denoising is also an ideal platform for testing natural image models and provides high-quality images for other computer vision tasks such as image registration, segmentation, and pattern recognition, etc.

For several decades, there emerge numerous image denoising methods [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11], and all of them focus mainly on dealing with additive white Gaussian noise (AWGN). In real world, the cameras will undertake high ISO settings for high-speed shots on actions, long exposure for low light on night shots, etc. Under these

situations, the noise is generated in a complex form and also been changed during the in-camera imaging pipeline [12, 13]. Therefore, the noise in real images are much more complex than Gaussian [13, 14]. It depends on camera series, brands, as well as the settings (ISO, shutter speed, and aperture, etc). The models designed for AWGN would become much less effective on real noisy images.

In the last decade, the methods of [15, 16, 17, 18, 19, 20, 13] are developed to deal with real noisy images. Almost all these methods employ a two-stage framework: estimating the parameters of the assumed noise model (usually Gaussian) and performing denoising with the help of the noise modeling and estimation in the first stage. However, the Gaussian assumption is inflexible in describing the complex noise on real noisy images [17]. Although the mixture of Gaussians (MoG) model is possible to approximate any noise distribution [21], estimating its parameters is time consuming via nonparametric Bayesian techniques [20]. To evaluate the performance of these methods on dealing with complex real noise, we apply these methods, with corresponding default parameters, on a real noisy image provided in [13]. The testing image is captured by a Nikon D800 camera when ISO is 3200. The "ground truth" image is also provided with which we can calculate objective measurements such as PSNR and SSIM [22]. The denoised images are listed in Figure 1, from which we can see that these methods either remove the noise or oversmooth the complex details in real noisy image.

The above mentioned methods can be categorized into external methods which learn priors from external images to recover noisy images, and internal ones which exploit priors of given images for denoising. The external priors in natural images are free of the high correlation between noise and signals in real noisy images, while the internal prior is adaptive to the image and can recover better the latent clean image. Combining the priors of external clean images and adaptively of internal testing images can naturally improve the performance of denoising methods, especially on real noisy images. Based on these observations, in this paper, we propose to employ the external patch group prior [10]

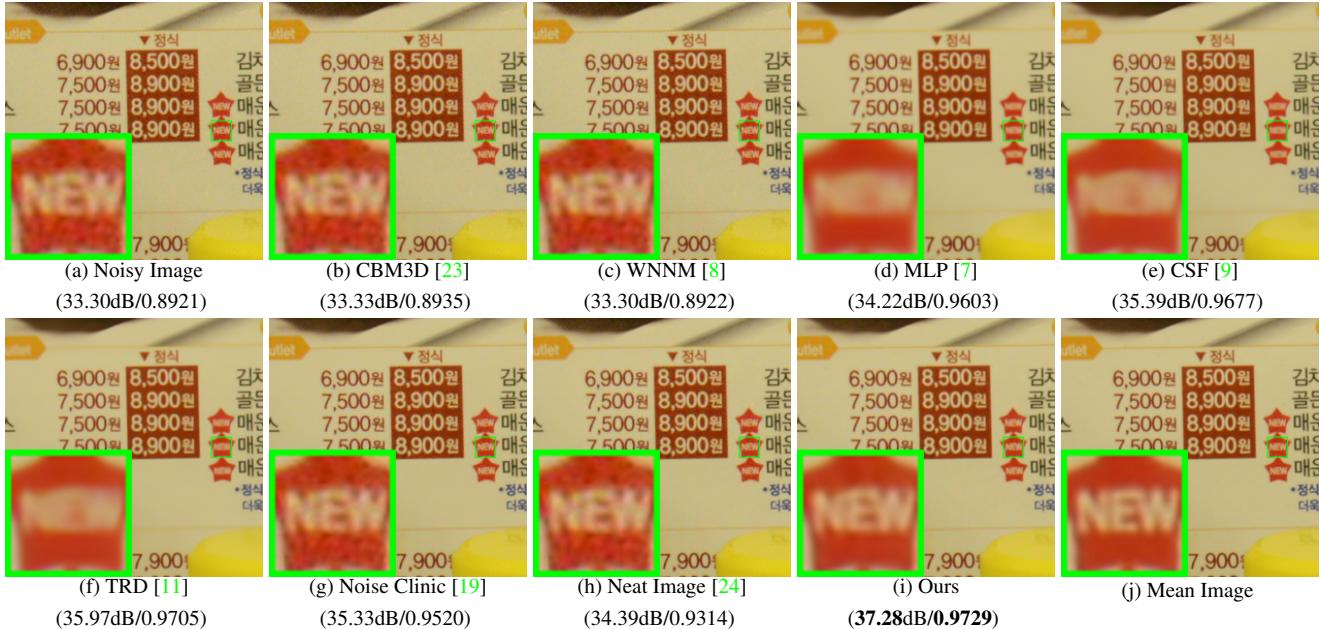


Figure 1. Denoised images of the real noisy image "Nikon D800 ISO 3200 A3" from [13] by different methods. The images are better viewed by zooming in on screen.

of natural clean images to guide the clustering of internal patch groups in given noisy image, and develop an external prior guided internal orthogonal dictionary learning (DL) algorithm for real image denoising. The internal orthogonal DL process includes two alternating stages: updating sparse coefficients and updating orthogonal dictionary. Both of the two stages have closed-form solutions. Hence, our internal DL process is very efficient for online internal denoising. Through comprehensive experiments on real noisy images captured by different cameras and settings, we demonstrate that the proposed method achieves better performance on real image denoising

1.1. Our Contributions

The contributions of this paper are summarized as follows:

- We propose a novel model to learn internal priors adaptive to given images. This model employs the external patch group (PG) prior learned from clean images to guide the internal PG prior learning of given images. The external prior benefits the internal learning on subspace selection and orthogonal dictionary learning.
- The proposed guided internal prior learning method is very efficient. The reason is that both the subspace selection and orthogonal dictionary learning have explicit solutions.
- For real image denoising problem, the proposed method achieves much better performance than other competing methods.

The rest of this paper will be summarized as follows: in Section 2, we briefly introduce the related work; in Section 3, we develop the proposed external prior guided internal prior learning model; in Section 4, we formulate the overall image denoising algorithm; in Section 4, we demonstrate extensive experiments on real image denoising probelm; in Section 5, we conclude our paper and give future work.

2. Related Work

2.1. Patch Group Prior of Natural Images

The Patch Group (PG) prior [10] is proposed to directly model the non-local self similar (NSS) property of natural images. The NSS property is commonly used in image restoration tasks [1, 4, 5, 8, 10]. The PG prior largely reduces the space of images to be modeled when compared to the patch prior [6]. In [10], only the PGs of clean natural images is utilized, while the PGs of noisy input images are ignored. In this paper, we make use of PGs both from external clean images and internal given real noisy image for better denoising performance.

2.2. Internal v.s. External Prior Learning

Learning priors to represent images has been successfully used in image modeling [3, 6, 10, 25, 33]. There are mainly two categories of prior learning methods: 1) External methods pre-learned priors (e.g., dictionaries) from a set of clean images, and the learned priors are used to recover the noisy images [6, 10]. 2) Internal methods directly learned priors from the given noisy image, and the image denoising is simultaneously done with the learning process

108 162
109 163
110 164
111 165
112 166
113 167
114 168
115 169
116 170
117 171
118 172
119 173
120 174
121 175
122 176
123 177
124 178
125 179
126 180
127 181
128 182
129 183
130 184
131 185
132 186
133 187
134 188
135 189
136 190
137 191
138 192
139 193
140 194
141 195
142 196
143 197
144 198
145 199
146 200
147 201
148 202
149 203
150 204
151 205
152 206
153 207
154 208
155 209
156 210
157 211
158 212
159 213
160 214
161 215

[3, 25, 33]. Both the two categories of methods have limitations. The external methods is not adaptive to the noisy image, while the internal methods ignores the information hidden in clean images. In this paper, our goal is to employ the external prior to guide the internal prior learning.

2.3. Real Image Denoising

In the last decade, there are many methods [15, 16, 17, 18, 19, 20, 13] proposed for real image denoising problem. In the seminar work of BLS-GSM [30] for real image denoising, Portilla et al. proposed to use scale mixture of Gaussian in overcomplete oriented pyramids to estimate the latent clean images. In [15], Portilla proposed to use a correlated Gaussian model for noise estimation of each wavelet subband. The work of Rabie [16] modeled the noisy pixels as outliers which are removed via Lorentzian robust estimator [31]. Liu et al. [17] proposed to use ‘noise level function’ to estimate the noise and then use Gaussian conditional random field to obtain the latent clean image. Gong et al. [18] models the noise by mixed ℓ_1 and ℓ_2 norms and remove the noise by sparsity prior in the wavelet transform domain. Later, Lebrun el al. proposed a multiscale denoising algorithm called ‘Noise Clinic’ [19]. This method generalizes the NL-Bayes model [32] to deal with blind noise and achieves state-of-the-art performance. Recently, Zhu et al. proposed a Bayesian model [20] which approximates and removes the noise via Low-Rank Mixture of Gaussians.

3. External Patch Group Prior Guided Internal Prior Learning

In this section, we formulate the framework of external patch group (PG) prior guided internal orthogonal dictionary learning. We first introduce the patch PG leaning on clean natural RGB images. Then we propose to employ the external PG prior to guide the internal clustering and orthogonal dictionary learning (DL). The orthogonal DL has alternative closed-form solutions in term of updating sparse coefficients and dictionary. Finally, we discuss the advantages of our proposed external PG prior guided internal orthogonal dictionary learning algorithm.

3.1. External Patch Group Prior Learning

Natural images often demonstrate repetitive local patterns, this nonlocal self-similarity (NSS) property is a key successful factor for many image denoising methods [1, 4, 5, 33, 8, 10]. In this section, we formulate the Patch Group prior learned on natural color images. Similar to [10], the patch group (PG) is defined as a group of similar patches to the local patch. The patch group mean is destracted, and hence different groups patches can share similar PGs. In this way, the space natural image patches to be modeled is largely reduced. In this work, each local patch

extracted from RGB images is of size $p \times p \times 3$. Then we search the M most similar patches $\{\mathbf{x}_m\}_{m=1}^M$ around each local patch through Euclidean distance, in a local window of size $W \times W$. The $\mathbf{x}_m \in \mathbb{R}^{3p^2 \times 1}$ is a patch vector formed by combining the 3 patch vectors (of size $p^2 \times 1$) in R, G, B channels. The mean vector of this PG is $\mu = \frac{1}{M} \sum_{m=1}^M \mathbf{x}_m$, and the group mean subtracted PG is defined as $\bar{\mathbf{X}} \triangleq \{\bar{\mathbf{x}}_m = \mathbf{x}_m - \mu\}, m = 1, \dots, M$. Assume we have extracted N PGs from a set of external natural images, and the n -th PG is defined as $\bar{\mathbf{X}}_n \triangleq \{\bar{\mathbf{x}}_{n,m}\}_{m=1}^M, n = 1, \dots, N$. We employ the Gaussian Mixture Model (GMM) to learn the external patch group based NSS prior. In this model, the likelihood of the n -th PG $\{\bar{\mathbf{X}}_n\}$ can be calculated as

$$P(\bar{\mathbf{X}}_n) = \sum_{k=1}^K \pi_k \prod_{m=1}^M \mathcal{N}(\bar{\mathbf{x}}_{n,m} | \mu_k, \Sigma_k), \quad (1)$$

where K is the number of Gaussians and the parameters π_k , μ_k , Σ_k are mixture weight, mean vector, and covariance matrix of the k -th Gaussian, respectively. By assuming that all the PGs are independently sampled, the overall objective log-likelihood function is

$$\ln \mathcal{L} = \sum_{n=1}^N \ln \left(\sum_{k=1}^K \pi_k \prod_{m=1}^M \mathcal{N}(\bar{\mathbf{x}}_{n,m} | \mu_k, \Sigma_k) \right). \quad (2)$$

We maximize the above objective function via EM algorithm [35] and finally obtain the GMM model with learned parameters as the learned external PG prior. Similar to [10], the mean vector of each cluster is natural zeros, i.e., $\mu_k = 0$.

3.2. External Prior Guided Internal Prior Learning

After the external patch group (PG) prior is learned, we can employ it to guide the internal PG prior learning for the given testing (real noisy) image. The guidance mainly comes from two aspects. One aspect is that the external prior can guide the internal noisy PGs to be assigned to most suitable Gaussians or subspaces. And for each subspace, the other aspect is to guide the orthogonal dictionary learning of internal noisy PGs.

3.2.1 Guided Internal Subspace Selection

Given a real noisy image, we extract the noisy PGs and corresponding mean vectors. Each mean substracted PG is defined as $\bar{\mathbf{Y}} \triangleq \{\bar{\mathbf{x}}_m\}_{m=1}^M$. Noted that, different from the external PGs, the mean vectors of the internal noisy PGs are saved for recovering. For adaptivity, we project the PG $\bar{\mathbf{Y}}$ into its most suitable Gaussian component (subspace) of the GMM learned on external PGs. The subspace most suitable for $\bar{\mathbf{Y}}$ is selected by firstly calculating the posterior probability of “ $\bar{\mathbf{Y}}$ belonging to the k th Gaussian component”:

$$P(k|\bar{\mathbf{Y}}) = \frac{\prod_{m=1}^M \mathcal{N}(\bar{\mathbf{y}}_m | \mathbf{0}, \Sigma_k)}{\sum_{l=1}^K \prod_{m=1}^M \mathcal{N}(\bar{\mathbf{y}}_m | \mathbf{0}, \Sigma_l)}, \quad (3)$$

and then choosing the component with the maximum A-posteriori (MAP) probability $\ln P(k|\bar{\mathbf{Y}})$.

324 **3.2.2 Guided Internal Orthogonal Dictionary Learn-**
 325 **ing**

326 Now each internal noisy PG has been assigned to its most
 327 suitable Gaussian component $\mathcal{N}(\mathbf{0}, \Sigma_k)$ (subspace). In order
 328 to recover the latent clean PG, we employ the sparse
 329 coding framework under which we need a suitable dictio-
 330 nary for coding. As we know, given a dictionary \mathbf{D} , its
 331 *mutual incoherence* $\mu(\mathbf{D})$ [36], which is difined as

$$\mu(\mathbf{D}) = \max_{i=j} \frac{|\mathbf{d}_i^T \mathbf{d}_j|}{\|\mathbf{d}_i\|_2 \|\mathbf{d}_j\|_2}, \quad (4)$$

332 is a measure of quality of dictionary. Since the *mutual incoher-
 333 ence* of a orthogonal dictionary is 0, its quality is often
 334 better than non-orthogonal dictionaries.

335 external subspace through covariance matrix via singular
 336 value decomposition (SVD)

$$\Sigma = \mathbf{U}_e \mathbf{S}_e \mathbf{V}_e^T. \quad (5)$$

337 can learn for each subspace an internal orthogonal dictio-
 338 nary adaptive to the noisy PGs. For notation simplicity, we
 339 ignore the index of subspace k .

$$\begin{aligned} & \min_{\mathbf{D}_i \in \mathbb{R}^{3p^2 \times r}, \mathbf{A} \in \mathbb{R}^{3p^2 \times MN}} \|\mathbf{Y} - [\mathbf{D}_e \mathbf{D}_i] \mathbf{A}\|_F^2 + \lambda \|\mathbf{A}\|_1 \\ & \text{s.t. } \mathbf{D}_i^T \mathbf{D}_i = \mathbf{I}_r, \mathbf{D}_e^T \mathbf{D}_i = \mathbf{0}, \end{aligned} \quad (6)$$

340 The singular vectors capture the statistical structures of NSS
 341 variations in natural images, while the singular values in \mathbf{S}
 342 represent the significance of these singular vectors. Fig. 4
 343 shows the singular vectors for one Gaussian component.

344 Similar to the K-SVD [3], we employ an alternating it-
 345 erative framework to solve the optimization problem 5. In
 346 fact, we initialize the orthogonal dictionary as $\mathbf{D}^{(0)}$ and for
 347 $t = 0, 1, \dots, T - 1$, alternatively do:

348 **Updating Sparse Coefficients:** given the initialization
 349 orthogonal dictioanry $\mathbf{D}_i^{(t)}$, the sparce coefficients $\mathbf{A}^{(t)}$ are
 350 obtained via solving

$$\mathbf{A}^{(t)} := \arg \min_{\mathbf{A} \in \mathbb{R}^{3p^2 \times MN}} \|\mathbf{Y} - [\mathbf{D}_e \mathbf{D}_i^{(t)}] \mathbf{A}\|_F^2 + \lambda \|\mathbf{A}\|_1. \quad (7)$$

351 This problem has closed-form solution by $\mathbf{A}^* =$
 352 $T_\lambda(\hat{\mathbf{D}}^T \mathbf{Y})$, where $T_\lambda(\mathbf{A}) = \text{sgn}(\mathbf{A}) \odot \max(\mathbf{A}, \lambda)$ is a soft-
 353 thresholding function.

354 **Updating Orthogonal Dictionary:** given the sparse co-
 355 efficients $\mathbf{A}^{(0)}$, the sparce coefficients $\mathbf{A}^{(t)}$ are obtained via
 356 solving

$$\begin{aligned} \mathbf{D}_i^{(t+1)} &:= \arg \min_{\mathbf{D}_i \in \mathbb{R}^{3p^2 \times r}} \|\mathbf{Y} - [\mathbf{D}_e \mathbf{D}_i] \mathbf{A}^{(t)}\|_F^2 \\ & \text{s.t. } \mathbf{D}_i^T \mathbf{D}_i = \mathbf{I}_r, \mathbf{D}_e^T \mathbf{D}_i = \mathbf{0}, \end{aligned} \quad (8)$$

357 Dividing the sparse coefficients $\mathbf{A} = [\mathbf{A}_e^T \mathbf{A}_i^T]^T$, where \mathbf{A}_e
 358 and \mathbf{A}_i denote the coefficients over external and internal

359 dictionary \mathbf{D}_e and \mathbf{D}_i . According to the Proposition 2.2
 360 in [34], the problem (7) has a closed-form solution $\mathbf{D}_i^* =$
 361 $\mathbf{U} \mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are the orthogonal matrices obtained
 362 by the following SVD

$$(\mathbf{I} - \mathbf{D}_e \mathbf{D}_e^T) \mathbf{Y} \mathbf{A}_i^T = \mathbf{U} \Sigma \mathbf{V}^T \quad (9)$$

363 With these solutions, the final obtained dictionary $\mathbf{D} =$
 364 $[\mathbf{D}_e \mathbf{D}_i]$ are orthogonal ictionary. This can be proved by
 365 the following equation

$$\mathbf{D}^T \mathbf{D} = \begin{pmatrix} \mathbf{D}_e^T \\ \mathbf{D}_i^T \end{pmatrix} (\mathbf{D}_e \mathbf{D}_i) = \begin{pmatrix} \mathbf{D}_e^T \mathbf{D}_e & \mathbf{D}_e^T \mathbf{D}_i \\ \mathbf{D}_i^T \mathbf{D}_e & \mathbf{D}_i^T \mathbf{D}_i \end{pmatrix} = \mathbf{I} \quad (10)$$

366 **3.3. Discussions on Learning External Priors and
 367 Internal Priors**

368 Until now, we have divided the noisy PGs into multiple
 369 internal subspaces. Here we take a deep analysis on how the
 370 external NSS prior guide the subspace learning of internal
 371 PGs. The help are at least threefold. Firstly, through MAP
 372 in (3), the external prior guides the noisy PGs to be clustered
 373 into the correct subspaces. If we cluster the noisy PGs in an
 374 automatical way, the subspaces we learned will be highly
 375 degraded by the signal dependent noise. Secondly, the guid-
 376 ance of external prior for internal clustering is more efficient
 377 than directly clustering the internal noisy PGs. It only needs
 378 to calculate the MAP probability via the equation (3) while
 379 the internal clustering via GMM is time-consuming on EM
 380 algorithm [35]. Thirdly, due to the correct guidance of ex-
 381 ternal prior, the strucutal decomposition via SVD of each
 382 subspace is more adaptive. This will bring better denois-
 383 ing performance than the methods only using the external
 384 information.

385 Through SVD, the PGs in each internal subspace can be
 386 divided into singular vectors and singular values. The sin-
 387 gular vectors are the basis of the corresponding subspace
 388 while the singular values reflect the importance of these ba-
 389 sis. The basis can be used as dictionary to code the noisy
 390 PGs. And the sigular values are adaptive parameters for in-
 391 ternal noisy PGs. We can compare the singular values of
 392 one internal subspace and the corresponding space of ex-
 393 ternal PGs. The result is shown in Figure ???. From which
 394 we can see that the noisy subspace often have higher val-
 395 ues than external space consisted of clean PGs. This gap is
 396 clearly made of the noise and can be used for image denois-
 397 ing in a natural way.

4. The Denoising Algorithm

4.1. Fast Patch Group Searching by Integral Image

398 The searching of patch groups in images is inefficient
 399 if we search non-local similar patches to each local patch.
 400 To speed up the searching process and make our proposed

method faster, we employ the technique of 'Summed Area Table' [37] for efficient PG searching. The SAT permits to evaluate the sum of pixel values in rectangular regions of the image with four operations, regardless of the region size. That is to say, we do not need do distance measure for each patch. It was first proposed under the name of summed area table[38]

4.2. Prior Weights for Sparse Coding

To remove the real noise, we employ the sparse coding framework. And in order to be adaptive to the input image, we employ the internal learned \mathbf{U} of each cluster as an adaptive dictioanry to represent the structural variations of the PGs in that cluster. Since the \mathbf{U} is orthonormal, its *mutual incoherence* is naturally 0 and therefore better than other redundant dictionaries.

$$\min_{\alpha} \|\bar{\mathbf{y}}_m - \mathbf{U}\alpha\|_2^2 + \sum_{i=1}^{3p^2} \lambda_i |\alpha_i|. \quad (11)$$

The i th entry of the regularization parameter λ_i

$$\lambda_i = \lambda / (\mathbf{S}_i + \varepsilon), \quad (12)$$

where ε is a small positive number to avoid dividing by zero. Since the dictionary \mathbf{U} is orthonormal, it is not difficult to find out that (5) has a closed-form solution (detailed derivation can be found in the supplementary material):

$$\hat{\alpha} = \text{sgn}(\mathbf{U}^T \bar{\mathbf{y}}_m) \odot \max(|\mathbf{U}^T \bar{\mathbf{y}}_m| - \Lambda, \mathbf{0}), \quad (13)$$

where $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_{3p^2}]$ is the vector of regularization parameter and $\text{sgn}(\bullet)$ is the sign function, \odot means element-wise multiplication, and $|\mathbf{U}^T \bar{\mathbf{y}}_m|$ is the absolute value of each entry of vector $|\mathbf{U}^T \bar{\mathbf{y}}_m|$. The closed-form solution makes our weighted sparse coding process very efficient.

4.3. The Overall Algorithm

With the solution $\hat{\alpha}$ in (9), the clean patch in a PG can be estimated as $\hat{\mathbf{x}}_m = \mathbf{D}\hat{\alpha} + \mu_y$. Then the clean image $\hat{\mathbf{x}}$ can be reconstructed by aggregating all the estimated PGs. In practice, we could perform the above denoising procedures for several iterations for better denoising outputs. In iteration t , we use the iterative regularization strategy [39] to add back to the recovered image $\hat{\mathbf{x}}^{(t-1)}$ some estimation residual in iteration $t-1$. The proposed denoising algorithm is summarized in Algorithm 1 (Alg. 1).

5. Experiments

In this section, we perform real image denoising experiments on three standard datasets. The first dataset is real noisy images with mean images as ground truths provided by [13], some samples are shown in Figure 3. The second dataset is provided by the website of Noise Clinic [19].

432	method faster, we employ the technique of 'Summed Area	486
433	Table' [37] for efficient PG searching. The SAT permits	487
434	to evaluate the sum of pixel values in rectangular regions	488
435	of the image with four operations, regardless of the region	489
436	size. That is to say, we do not need do distance measure	490
437	for each patch. It was first proposed under the name of summed	491
438	area table[38]	492
439		493
440		494
441		495
442		496
443		497
444		498
445		499
446		500
447		501
448		502
449		503
450		504
451		505
452		506
453		507
454		508
455		509
456		510
457		511
458		512
459		513
460		514
461		515
462		516
463		517
464		518
465		519
466		520
467		521
468		522
469		523
470		524
471		525
472		526
473		527
474		528
475		529
476		530
477		531
478		532
479		533
480		534
481		535
482		536
483		537
484		538
485		539

Alg. 1: External Prior Guided Internal Orthogonal Dictionary Learning for Denoising

Input: Noisy image \mathbf{y} , PG-GMM model

1. Initialization: $\hat{\mathbf{x}}^{(0)} = \mathbf{y}, \mathbf{y}^{(0)} = \mathbf{y}$;

for $t = 1 : IteNum$ do

 for each PG \mathbf{Y} do

 2. Calculate group mean μ_y and form PG $\bar{\mathbf{Y}}$;

 3. Gaussian component selection via (3);

 end for

 for each Internal Subspace do

 4. Internal Subspace Learning by (5);

 5. Recover each patch in all PGs via $\hat{\mathbf{x}}_m = \mathbf{D}\hat{\alpha} + \mu_y$;

 end for

 6. Aggregate the recovered PGs of all subspaces to form the recovered image $\hat{\mathbf{x}}^{(t)}$;

end for

Output: The recovered image $\hat{\mathbf{x}}^{(IteNum)}$.



Figure 2. Some testing images in the dataset [13].



Figure 3. Some cropped images of the dataset [13].

The third dataset is provided by the Commercial software Neat Image [24]. The second and third dataset do not have ground truth images.

5.1. Implementation Details

Our proposed method contains two stages, the external prior guided internal subspace learning stage and the adaptive denoising stage. In the learning stage, there are 4 parameters: the patch size p , the number of patches in a PG M , the window size W for PG searching and the number of clusters K . We set $p = 6$ (hence the patch size is $6 \times 6 \times 3$), $M = 10$, $W = 31$, $K = 32$. We extracted about 3.6 million PGs from the Kodak PhotoCD Dataset, which includes 24 high quality color images, to train the external prior via PG-GMM. In the denoising stage, the paramter $\lambda = 0.002$ is used to regularize the sparse term. The δ in iterative regularization is set as $\delta = 0.09$.

540 Table 1. Average PSNR(dB)/SSIM results of external, internal,
 541 and guided methods on 60 cropped real noisy images in [13].
 542

	Noisy	Offline	Online	Guided
PSNR	34.51	38.19	38.07	38.55
SSIM	0.8718	0.9663	0.9625	0.9675

5.2. Comparison on External and Internal methods

In this subsection, we compared the proposed external prior guided internal subspace learning model on real image denoising. The three methods are evaluated on the dataset provided in [13]. We calculate the PSNR, SSIM [22] and visual quality of these three methods. We also compare the speed. The PSNR and SSIM results on 60 cropped images from [13] are listed in Table 1. The images are cropped into size of 500×500 for better illustration. We also compare the three methods on visual quality in Figure 5.2. Compare the denoised images listed in Figure 5.2 and Figure 5.2, we can see that the Offline method is better at edges, smooth regions while the Online method is good at complex textures. The reason is two folds. Firstly, the Offline method is learned on clean images and hence is better at representing edges, structuals, and smooth area. The online method is influenced by the noise and hence some noise cannot be removed. Secondly, the Online method is better at recovering complex area sicne they could learn adaptive dictionaries for the specific area. The Offline method cannot recover the complex area since they did not learn the similar structures from the external natural clean images.

5.3. Comparison With other Competing Methods

We compare with previous state-of-the-art Gaussian noise removal methods such as BM3D [4], WNNM [8], MLP [7], CSF [9], and the recently proposed TRD [11]. We also compare with three competing real image denoising methods such as Noise Clinic, Neat Image, and the CC-Noise method proposed recently. The commercial software Neat Image [24] first estimates the parameters of noise via a large flat area and then filters the noise accordingly. All these methods need noise estimation which is vary hard to perform if there is no uniform regions are available in the testing image. The NeatImage will fail to perform automatical parameters settings if there is no uniform regions.¹

We the competing denoising methods from various research directions on two datasets. Both the two datasets comes from the [13]. The first dataset contains 17 images of size over 7000×5000 . Since this dataset contains repetitive contents across different images, we crop 60 small images of size 500×500 from these 17 images in [13]. The

¹To compare with CCNoise, we first transform the denoised images into double format.

PSNR and SSIM resluts are listed in Table 3. The number in red color and blue color means the best and second best results, respectively. From the Table 3, we can see that the external based method can already surpass largely the previous denoising methods. The improvement on PSNR over the second best method, i.e., TRD, is 0.44dB. The

5.4. Discussion on Parameter λ

The proposed method only has a key parameter, namely the regularization paramters λ . To demonstrate that the proposed method is robust to the variance of λ , we vary the parameter λ across a wide range and obtain the PSNR and SSIM results as a function of the parameter λ . The results is shown in Figure 8, from which we can see that the proposed method can achieve a PSNR (SSIM) over 38.5dB (0.9660) when λ varies from 0.0015 to 0.0025. This shows that the proposed method is indeed robust to the chosen of the paramter λ .

6. Conclusion and Future Work

In the future, we will evaluate the proposed method on other computer vision tasks such as single image super-resolution, photo-sketch synthesis, and cross-domain image recognition. Our proposed method can be improved if we use better training images, fine tune the parameters via cross-validation. We believe that our framework can be useful not just for real image denoising, but for image super-resolution, image cross-style synthesis, and recognition tasks. This will be our line of future work.

References

- [1] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. *CVPR*, pages 60–65, 2005. [1](#), [2](#), [3](#)
- [2] S. Roth and M. J. Black. Fields of Experts. *International Journal of Computer Vision*, 82(2):205–229, 2009. [1](#)
- [3] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *Image Processing, IEEE Transactions on*, 15(12):3736–3745, 2006. [1](#), [2](#), [3](#), [4](#)
- [4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *Image Processing, IEEE Transactions on*, 16(8):2080–2095, 2007. [1](#), [2](#), [3](#), [6](#)
- [5] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. *ICCV*, pages 2272–2279, 2009. [1](#), [2](#), [3](#)
- [6] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. *ICCV*, pages 479–486, 2011. [1](#), [2](#)
- [7] Harold C Burger, Christian J Schuler, and Stefan Harmeling. Image denoising: Can plain neural networks compete with bm3d? *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2392–2399, 2012. [1](#), [2](#), [6](#)

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755



Figure 4. Denoised images of the image "Nikon D600 ISO 3200 C1" by different methods. The images are better to be zoomed in on screen.

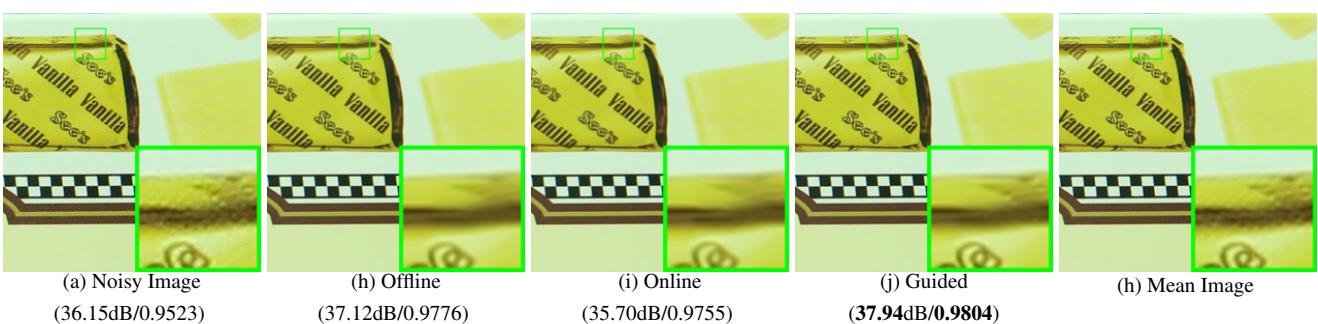


Figure 5. Denoised images of the image "Canon EOS 5D Mark3 ISO 3200 C1" by different methods. The images are better to be zoomed in on screen.

- [8] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. *CVPR*, pages 2862–2869, 2014. [1](#), [2](#), [3](#), [6](#)
- [9] U. Schmidt and S. Roth. Shrinkage fields for effective image restoration. *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 2774–2781, June 2014. [1](#), [2](#), [6](#)
- [10] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng. Patch group based nonlocal self-similarity prior learning for image denoising. *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 244–252, 2015. [1](#), [2](#), [3](#)
- [11] Yunjin Chen, Wei Yu, and Thomas Pock. On learning optimized reaction diffusion processes for effective image restoration. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5261–5269, 2015. [1](#), [2](#), [6](#)
- [12] S. J. Kim, H. T. Lin, Z. Lu, S. Ssstrunk, S. Lin, and M. S. Brown. A new in-camera imaging model for color computer vision and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(12):2289–2302, Dec 2012. [1](#)
- [13] 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. *Proc. Computer Vision and Pattern Recognition (CVPR)*, pages 1683–1691, 2016. [1](#), [2](#), [3](#), [5](#), [6](#), [8](#)
- [14] Glenn E Healey and Raghava Kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(3):267–276, 1994. [1](#)
- [15] J. Portilla. Full blind denoising through noise covariance estimation using gaussian scale mixtures in the wavelet domain. *Image Processing, 2004. ICIP '04. 2004 International Conference on*, 2:1217–1220, 2004. [1](#), [3](#)
- [16] Tamer Rabie. Robust estimation approach for blind denoising. *Image Processing, IEEE Transactions on*, 14(11):1755–1765, 2005. [1](#), [3](#)
- [17] C. Liu, R. Szeliski, S. Bing Kang, C. L. Zitnick, and W. T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):299–314, 2008. [1](#), [3](#)
- [18] Zheng Gong, Zuowei Shen, and Kim-Chuan Toh. Image restoration with mixed or unknown noises. *Multiscale Modeling & Simulation*, 12(2):458–487, 2014. [1](#), [3](#)
- [19] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image blind denoising. *Image Processing, IEEE Transactions on*, 24(10):3149–3161, 2015. [1](#), [2](#), [3](#), [5](#)
- [20] Fengyuan Zhu, Guangyong Chen, and Pheng-Ann Heng. From noise modeling to blind image denoising. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. [1](#), [3](#)
- [21] C. M. Bishop. *Pattern recognition and machine learning*. New York: Springer, 2006. [1](#)
- [22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to struc-

756 Table 2. Average PSNR(dB) results of different methods on 60 cropped real noisy images captured in [13].
757
758
759
760

	Noisy	CBM3D	WNNM	MLP	CSF	TRD	NI	NC	Guided	Guided2
PSNR	34.51	34.58	34.52	36.19	37.40	37.75	36.53	37.57	38.72	38.90
SSIM	0.8718	0.8748	0.8743	0.9470	0.9598	0.9617	0.9241	0.9514	0.9694	0.9702

761 Table 3. Average PSNR(dB) results of different methods on 15 cropped real noisy images used in [13].
762
763

Camera Settings	Noisy	CBM3D	WNNM	MLP	CSF	TRD	NI	NC	CC	Guided2
Canon 5D Mark III ISO = 3200	37.00	37.08	37.09	33.92	35.68	36.20	37.68	38.76	38.37	40.50
	33.88	33.94	33.93	33.24	34.03	34.35	34.87	35.69	35.37	37.22
	33.83	33.88	33.90	32.37	32.63	33.10	34.77	35.54	34.91	37.13
Nikon D600 ISO = 3200	33.28	33.33	33.34	31.93	31.78	32.28	34.12	35.57	34.98	35.34
	33.77	33.85	33.79	34.15	35.16	35.34	35.36	36.70	35.95	36.69
	34.93	35.02	34.95	37.89	39.98	40.51	38.68	39.28	41.15	39.17
Nikon D800 ISO = 1600	35.47	35.54	35.57	33.77	34.84	35.09	37.34	38.01	37.99	38.82
	35.71	35.79	35.77	35.89	38.42	38.65	38.57	39.05	40.36	40.98
	34.81	34.92	34.95	34.25	35.79	35.85	37.87	38.20	38.30	38.90
Nikon D800 ISO = 3200	33.26	33.34	33.31	37.42	38.36	38.56	36.95	38.07	39.01	38.69
	32.89	32.95	32.96	34.88	35.53	35.76	35.09	35.72	36.75	36.82
	32.91	32.98	32.96	38.54	40.05	40.59	36.91	36.76	39.06	38.80
Nikon D800 ISO = 6400	29.63	29.66	29.71	33.59	34.08	34.25	31.28	33.49	34.61	33.31
	29.97	30.01	29.98	31.55	32.13	32.38	31.38	32.79	33.21	33.18
	29.87	29.90	29.95	31.42	31.52	31.76	31.40	32.86	33.22	33.35
Average PSNR	33.41	33.48	33.48	34.32	35.33	35.65	35.49	36.43	36.88	37.26
Average SSIM	0.8483	0.8511	0.8512	0.9113	0.9250	0.9280	0.9126	0.9364	0.9481	0.9505

784 tural similarity. *IEEE Transactions on Image Processing*,
785 13(4):600–612, 2004. 1, 6787 [23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Color
788 image denoising via sparse 3d collaborative filtering with
789 grouping constraint in luminance-chrominance space. *IEEE*
790 *International Conference on Image Processing*, 1, 2007. 2791 [24] Neatlab ABSoft. Neat image. [https://ni.
792 neatvideo.com/home](https://ni.neatvideo.com/home). 2, 5, 6793 [25] G. Yu, G. Sapiro, and S. Mallat. Solving inverse problems
794 with piecewise linear estimators: From Gaussian mixture
795 models to structured sparsity. *IEEE Transactions on Image
796 Processing*, 21(5):2481–2499, 2012. 2, 3797 [26] Daniel Glasner, Shai Bagon, and Michal Irani. Super-
798 resolution from a single image. *ICCV*, 2009.800 [27] Maria Zontak, Inbar Mosseri, and Michal Irani. Separating
801 signal from noise using patch recurrence across scales. *Proceedings of the IEEE Conference on Computer Vision and
802 Pattern Recognition*, pages 1195–1202, 2013.804 [28] Tomer Michaeli and Michal Irani. Blind deblurring using
805 internal patch recurrence. *European Conference on Computer
806 Vision*, pages 783–798, 2014.807 [29] Y. Bahat and M. Irani. Blind dehazing using internal patch
808 recurrence. *IEEE International Conference on Computational Photography (ICCP)*, pages 1–9, May 2016.809 [30] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli.
Image denoising using scale mixtures of Gaussians in the
wavelet domain. *Image Processing, IEEE Transactions on*,
810 12(11):1338–1351, 2003. 3811 [31] Peter J Huber. *Robust statistics*. Springer, 2011. 3812 [32] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian
image denoising algorithm. *SIAM Journal on Imaging Sciences*,
813 6(3):1665–1688, 2013. 3814 [33] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally centralized
sparse representation for image restoration. *Image Processing, IEEE Transactions on*, 22(4):1620–1630, 2013. 2,
815 3816 [34] Chenglong Bao, Jian-Feng Cai, and Hui Ji. Fast sparsity-
based orthogonal dictionary learning for image restoration.
*Proceedings of the IEEE International Conference on Com-
puter Vision*, pages 3384–3391, 2013. 4817 [35] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum
likelihood from incomplete data via the EM algorithm. *Jour-
nal of the Royal Statistical Society. Series B (methodolog-
ical)*, pages 1–38, 1977. 3, 4818 [36] David L Donoho and Michael Elad. Optimally sparse rep-
resentation in general (nonorthogonal) dictionaries via 1 mini-
mization. *Proceedings of the National Academy of Sciences*,
819 100(5):2197–2202, 2003. 4

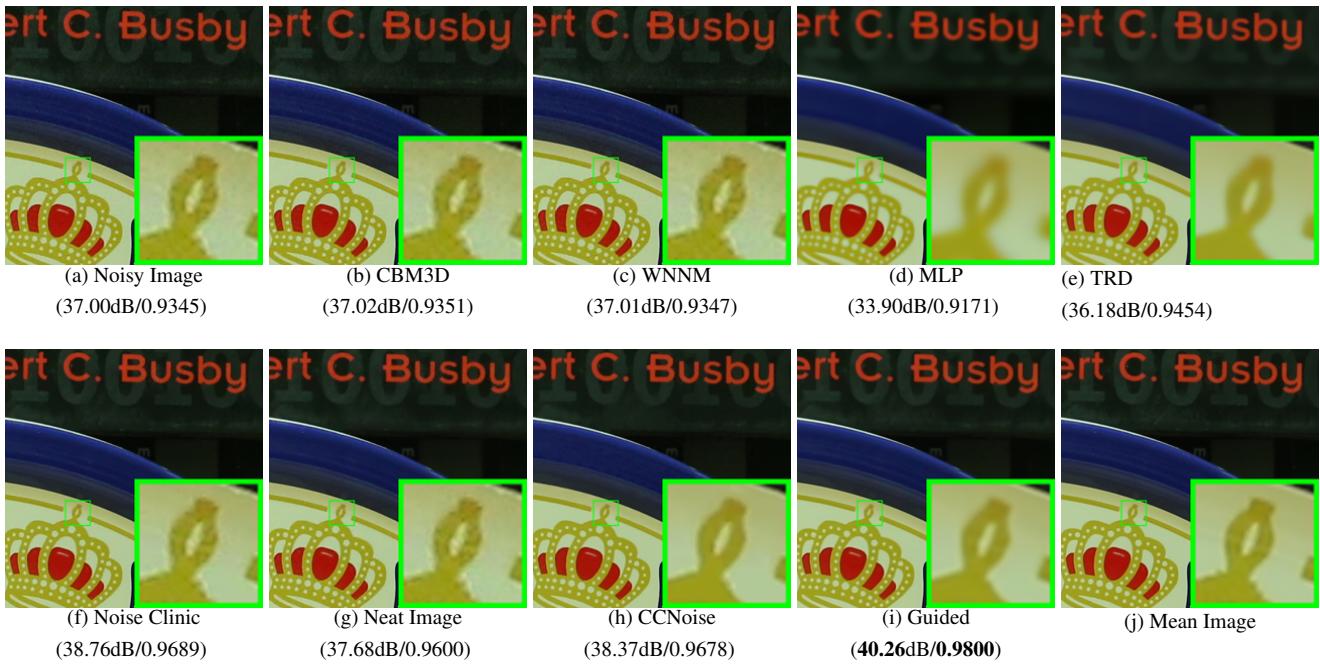


Figure 6. Denoised images of the image "Canon 5D Mark 3 ISO 3200 1" by different methods. The images are better to be zoomed in on screen.

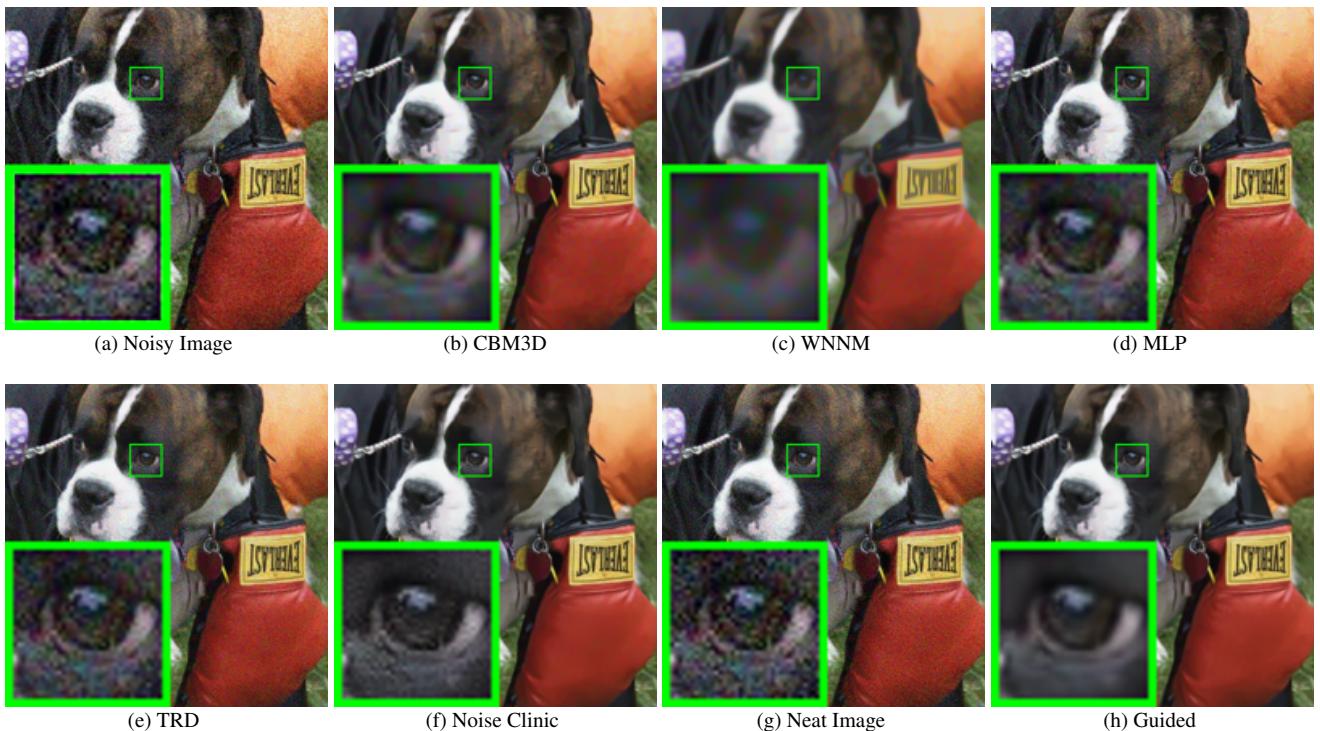


Figure 7. Denoised images of the image "5dmakr3iso32003" by different methods. The images are better to be zoomed in on screen.

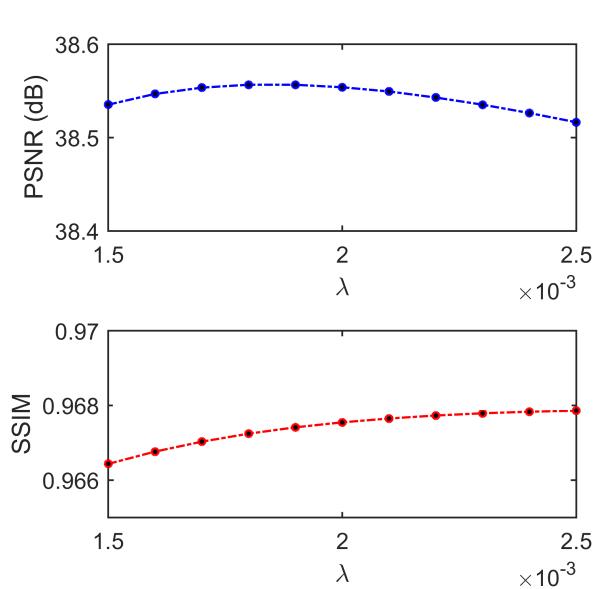
[37] Franklin C. Crow. Summed-area tables for texture mapping. *SIGGRAPH Comput. Graph.*, 18(3):207–212, January 1984.

5

[38] Paul Viola and Michael J Jones. Robust real-time face detec-

tion. *International journal of computer vision*, 57(2):137–154, 2004. 5

[39] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image



972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
Figure 8. The PSNR/SSIM results as a function of the parameter λ .

restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, 2005. 5

1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079