

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, 12], 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 [13, 14]. Therefore, the noise in real images are much more complex than Gaussian [14, 15]. 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 [16, 17, 18, 19, 20, 21, 22, 14] 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 [18]. Although the mixture of Gaussians (MoG) model is possible to approximate any noise distribution [23], estimating its parameters is time consuming via nonparametric Bayesian techniques [22]. 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 [14]. 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. 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 [11]

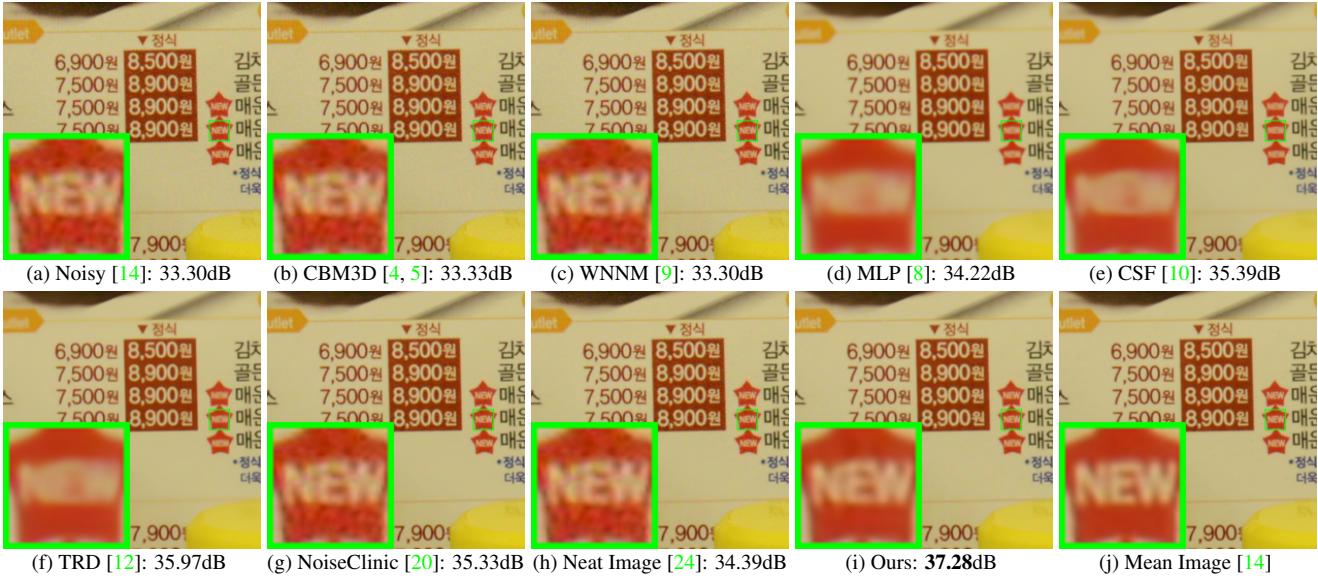


Figure 1. Denoised images of the real noisy image "Nikon D800 ISO 3200 A3" from [14] by different methods. The images are better viewed by zooming in on screen.

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 and formulate the overall image denoising algorithm; in Section 3, we demonstrate extensive experiments on real image denoising probelm; in Section 4, we conclude our paper and give future work.

## 2. Related Work

### 2.1. Patch Group Prior of Natural Images

The Patch Group (PG) prior [11] 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, 6, 9, 11]. The PG prior largely reduces the space of images to be modeled when compared to the patch prior [7]. In [11], 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, 7, 11, 25, 26]. 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 [7, 11]. 2) Internal methods directly learned priors from the given noisy image, and the image denoising is simultaneously done with the learning process [3, 25, 26]. Both the two categories of methods have lim-



Figure 2. Denoised images of the image "Nikon D600 ISO 3200 C1" by different methods. The images are better to be zoomed in on screen.

itations. 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, researchers proposed many methods [16, 17, 18, 19, 20, 21, 22, 14] for real image denoising problem. In [16], Portilla proposed to use scale mixture of Gaussian to estimate the noise covariance matrix and latent clean images in wavelet domain. The work of Rabie [17] modeled the noisy pixels as outliers which are removed via Lorentzian robust estimator [27]. Liu et al. [18] 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. [19] models the noise by mixed  $\ell_1$  and  $\ell_2$  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' [20, 21]. This method generalizes the NL-Bayes model [28] to deal with blind noise and achieves state-of-the-art performance. Recently, Zhu et al. proposed a Bayesian model [22] which approximates and removes the noise via Low-Rank Mixture of Gaussians.

## 3. External Patch Group Prior Guided Internal Prior Learning for Image Denoising

In this section, we formulate the framework of external patch group (PG) prior guided internal prior learning. We first introduce the external PG prior leaning on natural clean RGB images. Then we propose to employ the learned external PG prior to guide the internal prior (subspace selection and dictionary learning (DL)) learning of given degraded (such as noisy) images. Under the weighted sparse coding framework, the internal prior learning process has alternative closed-form solutions in term of updating sparse coefficients and orthogonal dictionary. Finally, we discuss in details how external prior learned from natural clean images guide the internal prior learning of given degraded (noisy) images.

### 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, 6, 26, 9, 11]. In this section, we formulate the Patch Group prior learned on natural color images. Similar to [11], the patch group (PG) is defined as a group of similar patches to the local patch. The patch group mean is distracted, 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  $\boldsymbol{\mu} = \frac{1}{M} \sum_{m=1}^M \mathbf{x}_m$ , and the group mean subtracted PG is defined as  $\bar{\mathbf{X}} \triangleq \{\bar{\mathbf{x}}_m = \mathbf{x}_m - \boldsymbol{\mu}\}, m = 1, \dots, M$ . 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

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

where  $K$  is the number of Gaussians and the parameters  $\pi_k$ ,  $\boldsymbol{\mu}_k$ ,  $\boldsymbol{\Sigma}_k$  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

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

We maximize the above objective function via EM algorithm [29] and finally obtain the GMM model with learned parameters. Similar to [11], the mean vector of each cluster is natural zeros, i.e.,  $\boldsymbol{\mu}_k = \mathbf{0}$ .

Now, we have clustered the PGs extracted from external clean images into  $K$  Gaussians or subspaces. To better

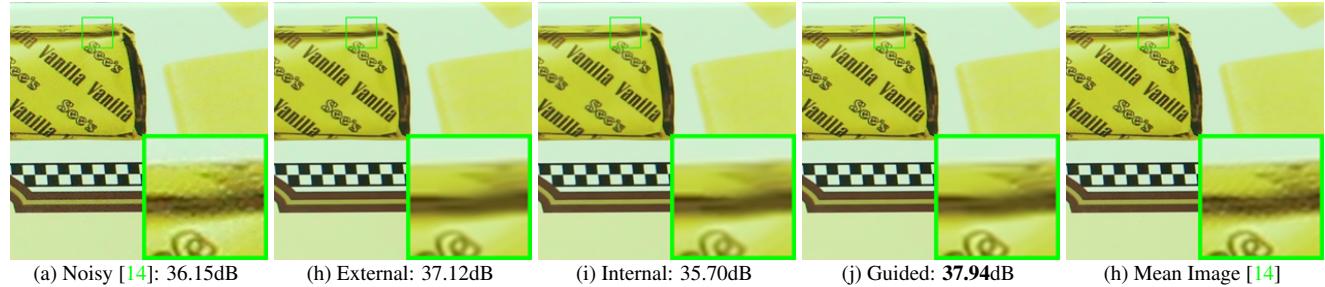


Figure 3. Denoised images of the image "Canon EOS 5D Mark3 ISO 3200 C1" by different methods. The images are better to be zoomed in on screen.

characterize each subspace, we perform singular value decomposition (SVD) on the covariance matrix:

$$\Sigma_k = \mathbf{U}_k \mathbf{S}_k \mathbf{U}_k^\top. \quad (3)$$

The singular vector matrices  $\{\mathbf{U}_k\}_{k=1}^K$  are employed as the external orthogonal dictionary to guide the internal dictionary learning. The singular values in the diagonal of  $\mathbf{S}_k$  reflect the significance of the singular vectors in  $\mathbf{U}_k$  and utilized as prior weights for weighted sparse coding which will be discussed in next section.

### 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, assume we can totally extract  $N$  local patches from it. Similar to the external prior learning stage, for the  $n$ -th local patch ( $n = 1, \dots, N$ ), we extract its  $M$  most similar patches around it to form a noisy PG denoted by  $\mathbf{Y}_n = \{\mathbf{y}_{n,1}, \dots, \mathbf{y}_{n,M}\}$ . Then the group mean of  $\mathbf{Y}_n$ , denoted by  $\boldsymbol{\mu}_n$ , is calculated and subtracted from each patch by  $\bar{\mathbf{y}}_{n,m} = \mathbf{y}_{n,m} - \boldsymbol{\mu}_n$ , leading to the mean subtracted PG  $\bar{\mathbf{Y}}_n \triangleq \{\bar{\mathbf{y}}_{n,m}\}_{m=1}^M$ . For adaptivity, we project the PG  $\bar{\mathbf{Y}}_n$  into its most suitable Gaussian component (subspace) of the GMM learned on external PGs. The subspace most suitable for  $\bar{\mathbf{Y}}_n$  is selected by firstly calculating the posterior probability of " $\bar{\mathbf{Y}}_n$  belonging to the  $k$ th Gaussian component":

$$P(k|\bar{\mathbf{Y}}_n) = \frac{\prod_{m=1}^M \mathcal{N}(\bar{\mathbf{y}}_{n,m}|\mathbf{0}, \Sigma_k)}{\sum_{l=1}^K \prod_{m=1}^M \mathcal{N}(\bar{\mathbf{y}}_{n,m}|\mathbf{0}, \Sigma_l)} \quad (4)$$

for  $k = 1, \dots, K$ , and then choosing the component with the maximum A-posteriori (MAP) probability  $\max_k P(k|\bar{\mathbf{Y}}_n)$ .

#### 3.2.2 Guided Internal Orthogonal Dictionary Learning

Assume we have assigned all internal noisy PGs  $\{\bar{\mathbf{Y}}_n\}_{n=1}^N$  to their corresponding most suitable Gaussians or subspaces in  $\{\mathcal{N}(\mathbf{0}, \Sigma_k)\}_{k=1}^K$ . For the  $k$ -th subspace, assume the noisy PGs assigned to it are  $\{\bar{\mathbf{Y}}_{k,n}\}_{n=1}^{N_k}$  such that  $\bar{\mathbf{Y}}_{k,n} = [\bar{\mathbf{y}}_{k,n,1}, \dots, \bar{\mathbf{y}}_{k,n,M}]$  and  $\sum_{k=1}^K N_k = N$ . We utilize the external orthogonal dictionary  $\mathbf{U}_k$  (Equ. (3)) to guide the learning of an orthogonal dictionary for adaptively characterizing the internal PGs in the  $k$ -th subspace. The learned dictionary defined as  $\mathbf{D}_k \triangleq [\mathbf{D}_{k,e} \ \mathbf{D}_{k,i}] \in \mathbb{R}^{3p^2 \times 3p^2}$  has two parts: the external part  $\mathbf{D}_{k,e} = \mathbf{U}_k(:, 1 : 3p^2 - r) \in \mathbb{R}^{3p^2 \times (3p^2 - r)}$  is directly obtained from the external dictionary  $\mathbf{U}_k$ , and the internal part  $\mathbf{D}_{k,i}$  is consisted of dictionary atoms adaptively learned from the internal noisy PGs  $\{\bar{\mathbf{Y}}_{k,n}\}_{n=1}^{N_k}$ . For notation simplicity, we ignore the subspace index  $k$  and denote the noisy PGs assigned to each subspace as  $\mathbf{Y} \triangleq \{\bar{\mathbf{Y}}_n\}_{n=1}^N = [\bar{\mathbf{y}}_{1,1}, \dots, \bar{\mathbf{y}}_{1,M}, \dots, \bar{\mathbf{y}}_{N,1}, \dots, \bar{\mathbf{y}}_{N,M}]$ . The learning is performed under the weighted sparse coding framework proposed as follows:

$$\begin{aligned} & \min_{\mathbf{D}_i, \{\alpha_{n,m}\}} \sum_{n=1}^N \sum_{m=1}^M (\|\bar{\mathbf{y}}_{n,m} - \mathbf{D}\alpha_{n,m}\|_2^2 + \sum_{j=1}^{3p^2} \lambda_j |\alpha_{n,m,j}|) \\ & \text{s.t. } \mathbf{D} = [\mathbf{D}_e \ \mathbf{D}_i], \ \mathbf{D}_i^\top \mathbf{D}_i = \mathbf{I}_r, \ \mathbf{D}_e^\top \mathbf{D}_i = \mathbf{0}, \end{aligned} \quad (5)$$

where  $\alpha_{n,m}$  is the sparse coefficient vector of the  $m$ -th patch  $\bar{\mathbf{y}}_{n,m}$  in the  $n$ -th PG  $\bar{\mathbf{Y}}_n$  and  $\alpha_{n,m,j}$  is the  $j$ -th element of  $\alpha_{n,m}$ .  $\lambda_j$  is the  $j$ -th regularization parameter defined as

$$\lambda_j = \lambda / (\sqrt{S_j} + \varepsilon). \quad (6)$$

We employ square roots of the singular values in  $\mathbf{S}$  (please refer to Equ. (3)) as external prior weights and add a small positive number  $\varepsilon$  to avoid zero denominator. Noted that  $\mathbf{D}_e = \emptyset$  if  $r = 3p^2$  and  $\mathbf{D}_e = \mathbf{U}_k$  if  $r = 0$ . The dictionary  $\mathbf{D} = [\mathbf{D}_e \ \mathbf{D}_i]$  is orthogonal by checking that:

$$\mathbf{D}^\top \mathbf{D} = \begin{bmatrix} \mathbf{D}_e^\top \\ \mathbf{D}_i^\top \end{bmatrix} [\mathbf{D}_e \ \mathbf{D}_i] = \begin{bmatrix} \mathbf{D}_e^\top \mathbf{D}_e & \mathbf{D}_e^\top \mathbf{D}_i \\ \mathbf{D}_i^\top \mathbf{D}_e & \mathbf{D}_i^\top \mathbf{D}_i \end{bmatrix} = \mathbf{I} \quad (7)$$

Similar to K-SVD [3], we employ an alternating iterative framework to solve the optimization problem (5). Specifically, we initialize the orthogonal dictionary as  $\mathbf{D}_{(0)} = \mathbf{U}_k$

432 and for  $t = 0, 1, \dots, T - 1$ , alternatively do:

433 **Updating Sparse Coefficient:** given the orthogonal dictionary  
 434  $\mathbf{D}_{(t)}$ , we update the sparse coefficient vector of the  $m$ -  
 435 th patch  $\bar{\mathbf{y}}_{n,m}$  in the  $n$ -th PG  $\bar{\mathbf{Y}}_n$  via solving

$$\alpha_{n,m}^{(t)} := \arg \min_{\alpha_{n,m}} \|\bar{\mathbf{y}}_{n,m} - \mathbf{D}_{(t)} \alpha_{n,m}\|_2^2 + \sum_{j=1}^{3p^2} \lambda_j |\alpha_{n,m,j}| \quad (8)$$

s.t.  $\mathbf{D}_{(t)} = [\mathbf{D}_e \mathbf{D}_i]$ ,  $\mathbf{D}_i^\top \mathbf{D}_i = \mathbf{I}_r$ ,  $\mathbf{D}_e^\top \mathbf{D}_i = \mathbf{0}$ ,

440 Since dictionary  $\mathbf{D}_{(t)} = [\mathbf{D}_e \mathbf{D}_i^{(t)}]$  is orthogonal, the problems (8) has a closed-form solution [11]

$$\alpha_{n,m}^{(t)} = \text{sgn}(\mathbf{D}_{(t)}^\top \bar{\mathbf{y}}_{n,m}) \odot \max(|\mathbf{D}_{(t)}^\top \bar{\mathbf{y}}_{n,m}| - \Lambda, \mathbf{0}), \quad (9)$$

441 where  $\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_{3p^2}]$  is the vector of regularization  
 442 parameter and  $\text{sgn}(\bullet)$  is the sign function,  $\odot$  means  
 443 element-wise multiplication.

444 **Updating Internal Dictionary:** given the sparse coefficient vectors  $\mathbf{A}^{(t)} = [\alpha_{1,1}^{(t)}, \dots, \alpha_{1,M}^{(t)}, \dots, \alpha_{N,1}^{(t)}, \dots, \alpha_{N,M}^{(t)}]$ ,  
 445 we update the internal orthogonal dictionary via solving

$$\begin{aligned} \mathbf{D}_i^{(t+1)} &:= \arg \min_{\mathbf{D}_i} \sum_{n=1}^N \sum_{m=1}^M (\|\bar{\mathbf{y}}_{n,m} - \mathbf{D}_i \alpha_{n,m}^{(t)}\|_2^2) \\ &= \arg \min_{\mathbf{D}_i} \|\mathbf{Y} - \mathbf{D} \mathbf{A}^{(t)}\|_F^2 \end{aligned} \quad (10)$$

$$\text{s.t. } \mathbf{D} = [\mathbf{D}_e \mathbf{D}_i], \mathbf{D}_i^\top \mathbf{D}_i = \mathbf{I}_r, \mathbf{D}_e^\top \mathbf{D}_i = \mathbf{0},$$

455 The sparse coefficient matrix  $\mathbf{A}^{(t)} = [(\mathbf{A}_e^{(t)})^\top, (\mathbf{A}_i^{(t)})^\top]^\top$   
 456 also has two parts: the external part  $\mathbf{A}_e^{(t)} \in \mathbb{R}^{(3p^2-r) \times NM}$  and the internal part  $\mathbf{A}_i^{(t)} \in \mathbb{R}^{r \times NM}$  denote the coefficients  
 457 over external dictionary  $\mathbf{D}_e$  and internal dictionary  $\mathbf{D}_i^{(t)}$ , respectively. According to the Theorem 4 in [30], the problem  
 458 (10) has a closed-form solution  $\mathbf{D}_i^{(t+1)} = \mathbf{U}_i \mathbf{V}_i^\top$ , where  
 459  $\mathbf{U}_i \in \mathbb{R}^{3p^2 \times 3p^2}$  and  $\mathbf{V}_i \in \mathbb{R}^{r \times r}$  are the orthogonal matrices obtained by the following SVD

$$(\mathbf{I} - \mathbf{D}_e \mathbf{D}_e^\top) \mathbf{Y} (\mathbf{A}_i^{(t)})^\top = \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^\top. \quad (11)$$

460 The orthogonality of internal dictionary  $\mathbf{D}_i^{(t+1)}$  can be  
 461 checked by  $(\mathbf{D}_i^{(t+1)})^\top (\mathbf{D}_i^{(t+1)}) = \mathbf{V}_i \mathbf{U}_i^\top \mathbf{U}_i \mathbf{V}_i^\top = \mathbf{I}_r$ .

### 3.3. Discussions

474 Here we take a detailed analysis on the guidance of the  
 475 external patch group (PG) prior for the internal noisy PGs  
 476 of given real noisy images. The guidance comes from at  
 477 least three aspects: 1) the external prior guides the internal  
 478 PGs to be clustered into suitable subspaces through MAP in  
 479 Equ. (4). The guided subspace selection is more efficient  
 480 than directly clustering the internal noisy PGs via k-means  
 481 or Gaussian Mixture Model (GMM). The reason is, by guid-  
 482 ance we only need calculate the probabilities via Equ. (4)  
 483 for all noisy PGs while by internal clustering via GMM we  
 484 have to perform time-consuming EM algorithm [29]. 2) the  
 485 external clean dictionary guides the learning of orthogonal

486 **Alg. 1:** External Patch Group (PG) Prior Guided Internal  
 487 PG Prior Learning for Image Denoising

488 **Input:** Noisy image  $\mathbf{y}$ , external PG prior GMM model

489 **Output:** The denoised image  $\hat{\mathbf{x}}$ .

490 **Initialization:**  $\hat{\mathbf{x}}^{(0)} = \mathbf{y}$ ;

491 **for**  $Ite = 1 : IteNum$  **do**

492   1. Extracting internal PGs from  $\hat{\mathbf{x}}^{(Ite-1)}$ ;

493   **for** each PG  $\mathbf{Y}_n$  **do**

494   2. Calculate group mean vector  $\mu_n$  and form  
 495   mean subtracted PG  $\bar{\mathbf{Y}}_n$ ;

496   3. Subspace selection via Equ. (4);

497   **end for**

498   **for** the PGs in each Subspace **do**

499   4. External PG prior Guided Internal Orthogonal  
 500   Dictionary Learning by solving (5);

501   5. Recover each patch in all PGs via Equ. (12);

502   **end for**

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

505   **end for**

506 dictionaries more adaptive for internal noisy images. The  
 507 learning process is very efficient because of closed-form  
 508 solutions. Besides, the learned orthogonal dictionary also  
 509 makes the denoising process very efficient under weighted  
 510 sparse coding framework. 3) the singular values learned by  
 511 SVD in Equ. (3) reflect the prior weights of the dictionary  
 512 atoms and can be used as adaptive parameters for real image  
 513 denoising.

### 3.4. The Denoising Algorithm

514 We evaluate the performance of the proposed framework  
 515 on denoising real noisy images. The denoising is simu-  
 516 lated done with the guided internal dictionary learn-  
 517 ing (DL) process. We ignore the index  $k \in \{1, \dots, K\}$  of  
 518 subspaces for notation simplicity. In the denoising stage,  
 519 for each subspace, the group mean vectors  $\{\mu_n\}_{n=1}^N$  of  
 520 corresponding mean subtracted noisy PGs  $\{\bar{\mathbf{Y}}_n\}_{n=1}^N$  are  
 521 saved for reconstruction. Until now, we obtain the sol-  
 522 tions of sparse coefficient vectors  $\{\hat{\alpha}_{n,m}^{(T-1)}\}$  in (9) for  
 523  $n = 1, \dots, N; m = 1, \dots, M$  and the orthogonal dictionary  
 524  $\mathbf{D}_{(T)} = [\mathbf{D}_e \mathbf{D}_i^{(T)}]$  in Equ. (10). Then the  $m$ -th latent clean  
 525 patch  $\hat{\mathbf{y}}_{n,m}$  in the  $n$ -th PG  $\bar{\mathbf{Y}}_n$  is recovered by

$$\hat{\mathbf{y}}_{n,m} = \mathbf{D}_{(T)} \hat{\alpha}_{n,m} + \mu_n, \quad (12)$$

526 where  $n = 1, \dots, N; m = 1, \dots, M$ . The latent clean im-  
 527 age  $\hat{\mathbf{x}}$  is reconstructed by aggregating all the estimated PGs.  
 528 Similar to [11], we perform the above denoising procedures  
 529 for several iterations for better denoising outputs. The pro-  
 530 posed denoising algorithm is summarized in Alg. 1.

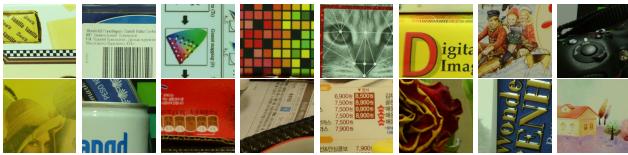


Figure 4. Some samples cropped from real noisy images of [14].

## 4. Experiments

In this section, we evaluate the performance of the proposed algorithm on real image denoising. To evaluation the effectiveness of the proposed framework of external prior guided internal prior learning, we compare it with the methods with only external prior or only internal prior. This will be discussed in Section 4.2. We also compare the proposed algorithm with other state-of-the-art denoising methods [4, 5, 8, 9, 10, 12, 14, 20, 21, 24].

### 4.1. The Testing Datasets

The comparisons are performed on three standard datasets, in which the images were captured by different types of cameras and camera settings under indoor and outdoor lighting conditions. The first dataset provided in [14] includes 17 real noisy images and corresponding mean images. The noisy images were collected under controlled indoor environment. Some samples can be found in [14]. For each image, the same scene was shot 500 times under the same camera and camera setting. The mean image of the 500 shots is roughly taken as the "ground truth", with which the PSNR can be computed for comparison. Since there are many similar contents in these 17 images, we cropped the 17 images into 60 smaller parts including different contents. Some samples are shown in Figure 4. The second dataset provided on [21] includes about 15 real noisy images collected by the authors of Noise Clinic [20]. The third dataset includes 18 color images provided by the website of Neat Image [24], a commercial software embeded in Photoshop CC and Corel PaintShop Pro X9 [24]. Different from the first dataset, the second and third datasets do not have ground truth images and hence the PSNR cannot be computed.

### 4.2. Implementation Details

Our proposed method contains two learning stages, the external prior learning stage and the external prior guided internal learning stage. In the external 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 Gaussians or subspaces  $K$  in Gaussian Mixture Model (GMM). 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

Table 1. Average PSNR(dB) results of External, Internal, and Guided methods on 60 cropped real noisy images in [14].

	Noisy	External	Internal	Guided
PSNR	34.51	38.21	38.07	<b>38.75</b>

prior via PG-GMM. In the denoising stage, the paramter  $\lambda = 0.002$  is used to regularize the sparse term. The  $\delta$  in iterative regularization is set as  $\delta = 0.09$ . Given a RGB image of size  $256 \times 256 \times 3$ , we can extract over 60 thousands local patches of size  $6 \times 6 \times 3$ . It is time-consuming to search patch groups (PG) for local patches one by one. To speed up the searching process, we employ the technique of 'Summed Area Table' [31] 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. Hence, we do not need do distance measure for each local patch.

### 4.3. 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 [14]. We compare the PSNR and visual quality of these three methods. We also compare the speed. The PSNR and SSIM results on 60 cropped images from [14] are listed in Table 1. The images are cropped into size of  $500 \times 500$  for better illustration. We also compare the three methods on visual quality in Figure 2. Compare the denoised images listed in Figure 2 and Figure 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 since 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.

### 4.4. Comparison With other Competing Methods

We compare with previous state-of-the-art Gaussian noise removal methods such as BM3D [4], WNNM [9], MLP [8], CSF [10], and the recently proposed TRD [12]. 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

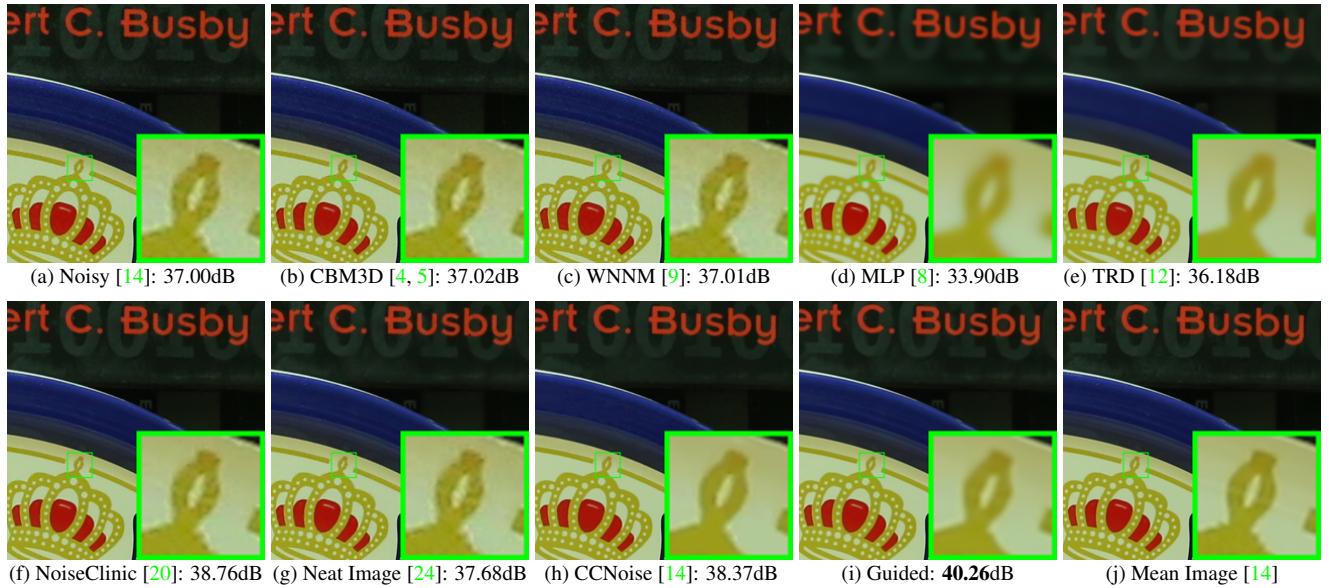


Figure 5. Denoised images of the image "Canon 5D Mark 3 ISO 3200 1" by different methods. The images are better to be zoomed in on screen.

Table 2. Average PSNR(dB) results of different methods on 60 cropped real noisy images captured in [14].

	Noisy	CBM3D	WNNM	MLP	CSF	TRD	NI	NC	Guided
PSNR	34.51	34.58	34.52	36.19	37.40	37.75	36.53	37.57	38.75

676 perform if there is no uniform regions are available in the  
677 testing image. The NeatImage will fail to perform automatical  
678 parameters settings if there is no uniform regions.<sup>1</sup>

679 We the competing denoising methods from various re-  
680 search directions on two datasets. Both the two datasets  
681 comes from the [14]. The first dataset contains 17 images  
682 of size over  $7000 \times 5000$ . Since this dataset contains repet-  
683 itive contents across different images, we crop 60 small im-  
684 ages of size  $500 \times 500$  from these 17 images in [14]. The  
685 PSNR and SSIM resluts are listed in Table 3. The number  
686 in red color and blue color means the best and second best  
687 results, respectively. From the Table 3, we can see that the  
688 external based method can already surpass largely the pre-  
689 vious denoising methods. The improvement on PSNR over  
690 the second best method, i.e., TRD, is 0.44dB. The

#### 691 4.5. Discussion on Parameter $\lambda$

693 The proposed method only has a key parameter, namely  
694 the regularization paramters  $\lambda$ . To demonstrate that the pro-  
695 posed method is robust to the variance of  $\lambda$ , we vary the  
696 parameter  $\lambda$  across a wide range and obtain the PSNR and  
697 SSIM results as a function of the parameter  $\lambda$ . The re-  
698 sults is shown in Figure ??, from which we can see that the  
699 proposed method can achieve a PSNR (SSIM) over 38.5dB  
700

701 <sup>1</sup>To compare with CCNoise, we first transform the denoised images  
702 into double format.

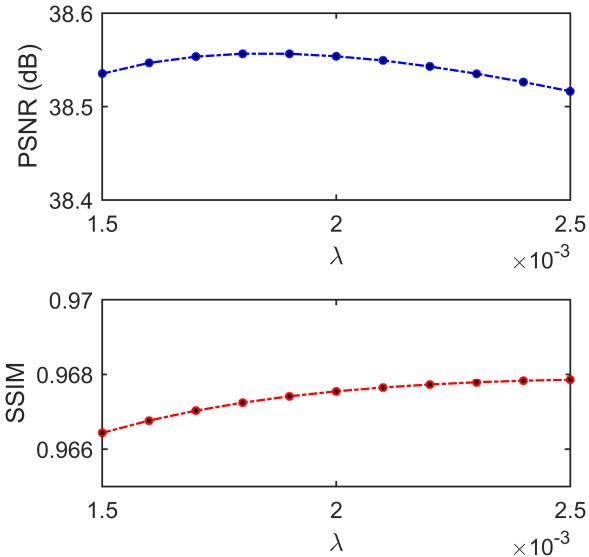


Figure 6. The PSNR results as a function of the parameter  $\lambda$ .

(0.9660) when  $\lambda$  varies from 0.0015 to 0.0025. This shows that the proposed method is indeed robust to the chosen of the paramter  $\lambda$ .

Table 3. Average PSNR(dB) results of different methods on 15 cropped real noisy images used in [14].

Camera Settings	Noisy	CBM3D	WNNM	MLP	CSF	TRD	NI	NC	CC	Guided
Canon 5D Mark III ISO = 3200	37.00 33.88 33.83	37.08 33.94 33.88	37.09 33.93 33.90	33.92 33.24 32.37	35.68 34.03 32.63	36.20 34.35 33.10	37.68 34.87 34.77	38.76 35.69 35.54	38.37 35.37 34.91	40.50 37.05 36.11
Nikon D600 ISO = 3200	33.28 33.77 34.93	33.33 33.85 35.02	33.34 33.79 34.95	31.93 34.15 37.89	31.78 35.16 39.98	32.28 35.34 40.51	34.12 35.36 38.68	35.57 36.70 41.15	34.98 35.95 39.28	34.88 36.31 39.23
Nikon D800 ISO = 1600	35.47 35.71 34.81	35.54 35.79 34.92	35.57 35.77 34.95	33.77 35.89 34.25	34.84 38.42 35.79	35.09 38.65 35.85	37.34 38.57 37.87	38.01 39.05 38.20	37.99 40.36 38.30	38.40 40.92 38.97
Nikon D800 ISO = 3200	33.26 32.89 32.91	33.34 32.95 32.98	33.31 32.96 32.96	37.42 34.88 38.54	38.36 35.53 40.05	38.56 35.76 40.59	36.95 35.09 36.91	38.07 35.72 36.76	39.01 36.75 39.06	38.66 37.07 38.52
Nikon D800 ISO = 6400	29.63 29.97 29.87	29.66 30.01 29.90	29.71 29.98 29.95	33.59 31.55 31.42	34.08 32.13 31.52	34.25 32.38 31.76	31.28 31.38 31.40	33.49 32.79 32.86	34.61 33.21 33.22	33.76 33.43 33.58
Average	33.41	33.48	33.48	34.32	35.33	35.65	35.49	36.43	36.88	37.16

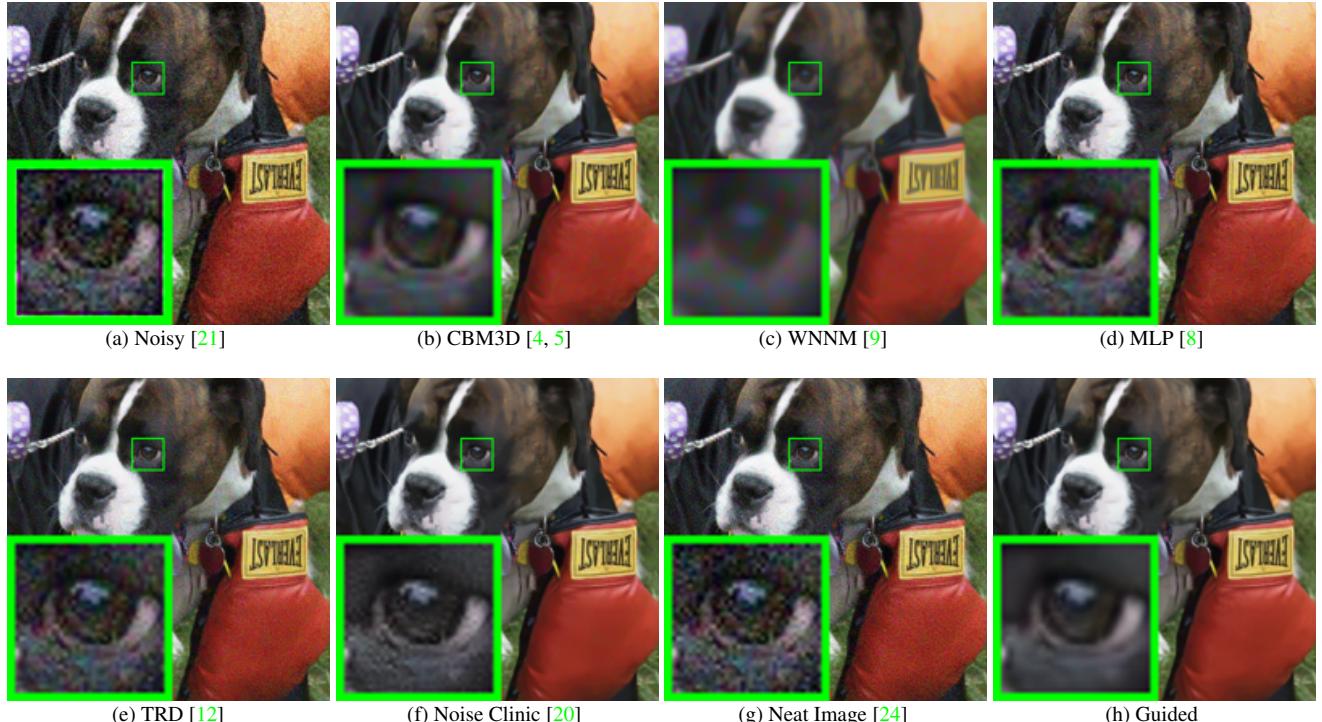


Figure 7. Denoised images of the image "5dmak3iso32003" by different methods. The images are better to be zoomed in on screen.

## 5. 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

- 864 [2] S. Roth and M. J. Black. Fields of Experts. *International  
865 Journal of Computer Vision*, 82(2):205–229, 2009. 1
- 866 [3] M. Elad and M. Aharon. Image denoising via sparse and  
867 redundant representations over learned dictionaries. *Image  
868 Processing, IEEE Transactions on*, 15(12):3736–3745,  
869 2006. 1, 2, 4
- 870 [4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image  
871 denoising by sparse 3-D transform-domain collaborative  
872 filtering. *Image Processing, IEEE Transactions on*,  
873 16(8):2080–2095, 2007. 1, 2, 3, 6, 7, 8
- 874 [5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Color  
875 image denoising via sparse 3d collaborative filtering with  
876 grouping constraint in luminance-chrominance space. *IEEE  
877 International Conference on Image Processing*, 1, 2007. 1,  
878 2, 6, 7, 8
- 879 [6] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman.  
880 Non-local sparse models for image restoration. *ICCV*, pages  
881 2272–2279, 2009. 1, 2, 3
- 882 [7] D. Zoran and Y. Weiss. From learning models of natural  
883 image patches to whole image restoration. *ICCV*, pages 479–  
884 486, 2011. 1, 2
- 885 [8] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising:  
886 Can plain neural networks compete with bm3d? *Computer  
887 Vision and Pattern Recognition (CVPR), 2012 IEEE  
888 Conference on*, pages 2392–2399, 2012. 1, 2, 6, 7, 8
- 889 [9] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear  
890 norm minimization with application to image denoising.  
891 *CVPR*, pages 2862–2869, 2014. 1, 2, 3, 6, 7, 8
- 892 [10] U. Schmidt and S. Roth. Shrinkage fields for effective im-  
893 age restoration. *Computer Vision and Pattern Recognition  
894 (CVPR), 2014 IEEE Conference on*, pages 2774–2781, June  
895 2014. 1, 2, 6
- 896 [11] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng. Patch  
897 group based nonlocal self-similarity prior learning for image  
898 denoising. *2015 IEEE International Conference on Computer  
899 Vision (ICCV)*, pages 244–252, 2015. 1, 2, 3, 5
- 900 [12] Y. Chen, W. Yu, and T. Pock. On learning optimized reaction  
901 diffusion processes for effective image restoration. *Proceed-  
902 ings of the IEEE Conference on Computer Vision and Pattern  
903 Recognition*, pages 5261–5269, 2015. 1, 2, 6, 7, 8
- 904 [13] S. J. Kim, H. T. Lin, Z. Lu, S. Ssstrunk, S. Lin, and M. S.  
905 Brown. A new in-camera imaging model for color computer  
906 vision and its application. *IEEE Transactions on Pattern  
907 Analysis and Machine Intelligence*, 34(12):2289–2302, Dec  
908 2012. 1
- 909 [14] S. Nam, Y. Hwang, Y. Matsushita, and S. J. Kim. A holis-  
910 tic approach to cross-channel image noise modeling and its  
911 application to image denoising. *Proc. Computer Vision and  
912 Pattern Recognition (CVPR)*, pages 1683–1691, 2016. 1, 2,  
913 3, 4, 6, 7, 8
- 914 [15] G. E Healey and R. Kondepudy. Radiometric ccd camera cal-  
915 ibration and noise estimation. *IEEE Transactions on Pattern  
916 Analysis and Machine Intelligence*, 16(3):267–276, 1994. 1
- 917 [16] J. Portilla. Full blind denoising through noise covariance  
918 estimation using gaussian scale mixtures in the wavelet do-  
919 main. *Image Processing, 2004. ICIP '04. 2004 International  
920 Conference on*, 2:1217–1220, 2004. 1, 3
- 921 [17] T. Rabie. Robust estimation approach for blind denois-  
922 ing. *Image Processing, IEEE Transactions on*, 14(11):1755–  
923 1765, 2005. 1, 3
- 924 [18] C. Liu, R. Szeliski, S. Bing Kang, C. L. Zitnick, and W. T.  
925 Freeman. Automatic estimation and removal of noise from  
926 a single image. *IEEE Transactions on Pattern Analysis and  
927 Machine Intelligence*, 30(2):299–314, 2008. 1, 3
- 928 [19] Z. Gong, Z. Shen, and K.-C. Toh. Image restoration with  
929 mixed or unknown noises. *Multiscale Modeling & Simula-  
930 tion*, 12(2):458–487, 2014. 1, 3
- 931 [20] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image  
932 blind denoising. *Image Processing, IEEE Transactions on*,  
933 24(10):3149–3161, 2015. 1, 2, 3, 6, 7, 8
- 934 [21] M. Lebrun, M. Colom, and J. M. Morel. The noise clinic:  
935 a blind image denoising algorithm. [http://www.ipol.  
936 im/pub/art/2015/125/](http://www.ipol.im/pub/art/2015/125/). Accessed 01 28, 2015. 1, 3,  
937 6, 8
- 938 [22] F. Zhu, G. Chen, and P.-A. Heng. From noise modeling to  
939 blind image denoising. *The IEEE Conference on Computer  
940 Vision and Pattern Recognition (CVPR)*, June 2016. 1, 3
- 941 [23] C. M. Bishop. *Pattern recognition and machine learning*.  
942 New York: Springer, 2006. 1
- 943 [24] Neatlab ABSsoft. Neat image. [https://ni.  
944 neatvideo.com/home](https://ni.neatvideo.com/home). 2, 6, 7, 8
- 945 [25] G. Yu, G. Sapiro, and S. Mallat. Solving inverse problems  
946 with piecewise linear estimators: From Gaussian mixture  
947 models to structured sparsity. *IEEE Transactions on Image  
948 Processing*, 21(5):2481–2499, 2012. 2
- 949 [26] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally central-  
950 ized sparse representation for image restoration. *Image Pro-  
951 cessing, IEEE Transactions on*, 22(4):1620–1630, 2013. 2,  
952 3
- 953 [27] P. J. Huber. *Robust statistics*. Springer, 2011. 3
- 954 [28] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian  
955 image denoising algorithm. *SIAM Journal on Imaging Sci-  
956 ences*, 6(3):1665–1688, 2013. 3
- 957 [29] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum  
958 likelihood from incomplete data via the EM algorithm. *Jour-  
959 nal of the Royal Statistical Society. Series B (methodologi-  
960 cal)*, pages 1–38, 1977. 3, 5
- 961 [30] H. Zou, T. Hastie, and R. Tibshirani. Sparse principal com-  
962 ponent analysis. *Journal of computational and graphical  
963 statistics*, 15(2):265–286, 2006. 5
- 964 [31] F. C. Crow. Summed-area tables for texture mapping. *SIG-  
965 GRAPH Comput. Graph.*, 18(3):207–212, January 1984. 6
- 966 [918] 967 [919] 968 [920] 969 [921] 970 [922] 971