

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]

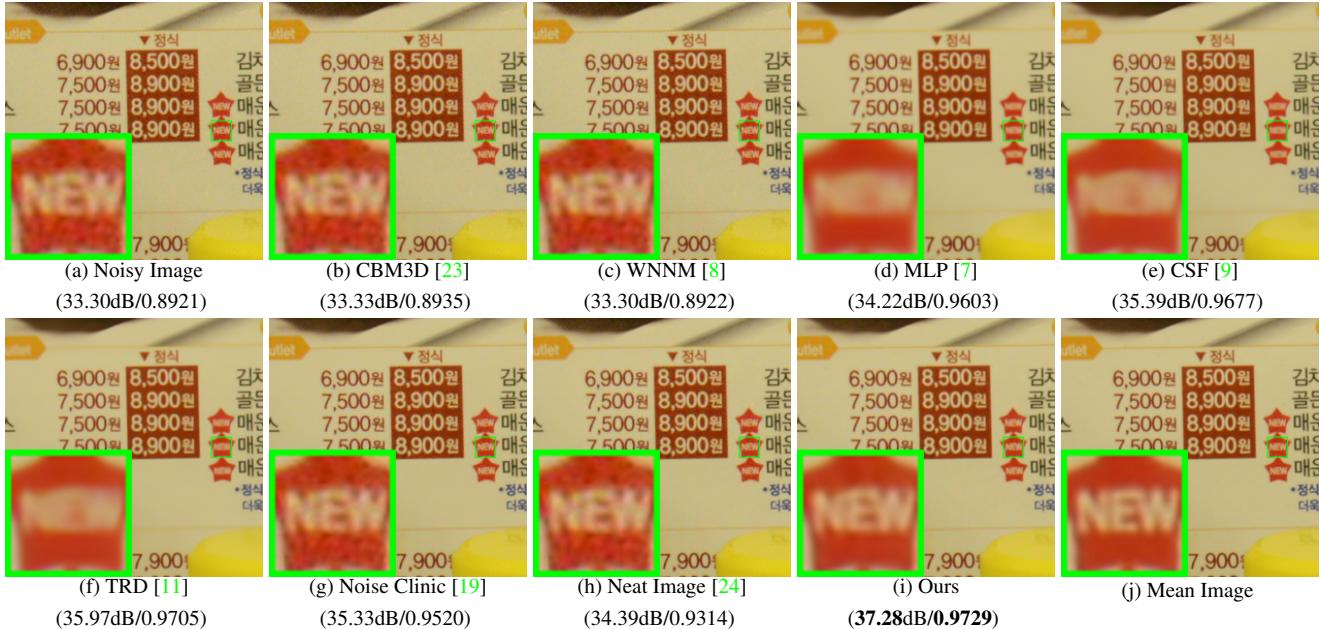


Figure 1. Denoised images of the real noisy image "Nikon D800 ISO 3200 A3" from [13] by different methods. The images are better viewed by zooming in on screen.

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 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 [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, 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 [6, 10]. 2) Internal methods directly learned priors from the given noisy image, and the image denoising is simultaneously done with the learning process



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.

[3, 25, 26]. 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 [27] 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 [28]. 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  $\ell_1$  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' [19]. This method generalizes the NL-Bayes model [29] 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 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 co-

efficients 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, 5, 26, 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 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

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

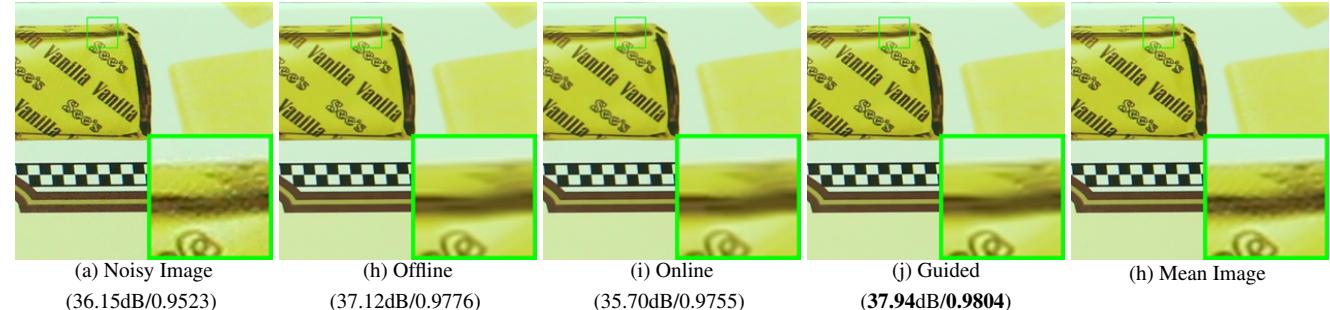


Figure 3. Denoised images of the image "Canon EOS 5D Mark3 ISO 3200 C1" by different methods. The images are better to be zoomed in on screen.

We maximize the above objective function via EM algorithm [30] and finally obtain the GMM model with learned parameters. Similar to [10], the mean vector of each cluster is natural zeros, i.e.,  $\mu_k = \mathbf{0}$ .

Now, we have clustered the PGs extracted from external clean images into  $K$  Gaussians or subspaces. To better 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 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, leading to the mean subtracted PG  $\bar{\mathbf{Y}}_n$ . Each mean subtracted PG is defined as  $\bar{\mathbf{Y}}_n \triangleq \{\bar{\mathbf{y}}_{n,m}\}_{m=1}^M$  where  $\bar{\mathbf{y}}_{n,m} = \mathbf{y}_{n,m} - \boldsymbol{\mu}_n$ . 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":

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

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

#### 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  $\{\mathcal{N}(\mathbf{0}, \Sigma_k)\}_{k=1}^K$ . Assume  $\{\bar{\mathbf{Y}}_{k_n}\}_{n=1}^{N_k}$  are the noisy PGs assigned to the  $k$ -th subspace, such that  $\sum_{k=1}^K N_k = N$  and  $\sum_{k=1}^K \sum_{n=1}^{N_k} k_n = NM$ , we have that  $\bar{\mathbf{Y}}_{k_n} = [\mathbf{y}_{k_n,1}, \dots, \mathbf{y}_{k_n,M}]$ . We consider to utilize the external orthogonal dictionary  $\mathbf{U}_k$  to guide the learning of an orthogonal dictionary  $\mathbf{D}_k := [\mathbf{D}_{k,e} \ \mathbf{D}_{k,i}] \in \mathbb{R}^{3p^2 \times 3p^2}$ . This dictionary 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 in Equ. (3), and the internal part  $\mathbf{D}_{k,i}$  is consisted of dictionary atoms adaptively learned from the internal noisy PGs  $\mathbf{Y}_k$ . The learning and denoising are simultaneously performed under the proposed reweighted sparse coding framework as f(for notation simplicity, we ignore the subspace index  $k$ ):

$$\min_{\mathbf{D}_i, \{\boldsymbol{\alpha}_n\}} \sum_{n=1}^N (\|\mathbf{y}_n - \mathbf{D}\boldsymbol{\alpha}_n\|_2^2 + \sum_{j=1}^{3p^2} \lambda_j |\boldsymbol{\alpha}_{n,j}|) \quad (5)$$

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

where  $\lambda_j$  is the  $j$ -th regularization parameter defined as

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

We employ square roots of the singular values in  $\mathbf{S}$  (please refer 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  $\mathbf{D}_{(0)} = \mathbf{U}_k$  and for  $t = 0, 1, \dots, T - 1$ , alternatively do:

**Updating Sparse Coefficients:** given the orthogonal dictionary  $\mathbf{D}_{(t)}$ , we update the sparse coefficients via solving

$$\boldsymbol{\alpha}_n^{(t)} := \arg \min_{\boldsymbol{\alpha}_n \in \mathbb{R}^{3p^2 \times 1}} \|\mathbf{y}_n - \mathbf{D}_{(t)} \boldsymbol{\alpha}_n\|_2^2 + \sum_{j=1}^{3p^2} \lambda_j |\alpha_{n,j}| \quad (8)$$

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

Since dictionary  $\mathbf{D}_{(t)} = [\mathbf{D}_e \mathbf{D}_i^{(t)}]$  is orthogonal, the problem (8) has a closed-form solution [10]

$$\hat{\boldsymbol{\alpha}}_n^{(t)} = \text{sgn}(\mathbf{D}_{(t)}^\top \mathbf{y}_n) \odot \max(|\mathbf{D}_{(t)}^\top \mathbf{y}_n| - \Lambda, 0), \quad (9)$$

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.

**Updating Orthogonal Dictionary:** given the sparse coefficients  $\mathbf{A}^{(t)} = [\boldsymbol{\alpha}_1^{(t)}, \dots, \boldsymbol{\alpha}_N^{(t)}]$ , we update the internal orthogonal dictionary via solving

$$\begin{aligned} \mathbf{D}_i^{(t+1)} &:= \arg \min_{\mathbf{D}_i} \sum_{n=1}^N (\|\mathbf{y}_n - \mathbf{D} \boldsymbol{\alpha}_n^{(t)}\|_2^2) \\ &= \arg \min_{\mathbf{D}_i} \|\mathbf{Y} - \mathbf{D} \mathbf{A}^{(t)}\|_F^2 \end{aligned} \quad (10)$$

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

Here we ignore the index  $t$  for notation simplicity. The sparse coefficient matrix  $\mathbf{A} = [\mathbf{A}_e^\top \mathbf{A}_i^\top]^\top$  also has two parts: the external part  $\mathbf{A}_e$  and the internal part  $\mathbf{A}_i$  denote the coefficients over external dictionary  $\mathbf{D}_e$  and internal dictionary  $\mathbf{D}_i$ , respectively. According to the Proposition 2.2 in [33], the problem (9) has a closed-form solution  $\mathbf{D}_i^* = \mathbf{U}_i \mathbf{V}_i^\top$ , where  $\mathbf{U}_i$  and  $\mathbf{V}_i$  are the orthogonal matrices obtained by the following SVD

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

### 3.3. Discussions

Here we take a detailed analysis on the guidance of the external patch group (PG) prior for the internal noisy PGs of given real noisy images. The guidance comes from at least three aspects: 1) the external prior guides the internal PGs to be clustered into suitable subspaces through MAP in Equ. (5). The guided subspace clustering is more efficient than directly clustering the internal noisy PGs via k-means or Gaussian Mixture Model (GMM). The reason is, by guidance we only need calculate the probabilities via Equ. (5) for all noisy PGs while by internal clustering via GMM we need perform EM algorithm [30]. 2) the external dictionary guides the internal dictionary learning, and the obtained dictionary consisted of the two parts is orthogonal and more

<b>Alg. 1:</b> External PG Prior Guided Internal PG Prior Learning for Real Image Denoising	486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
<b>Input:</b> Noisy image $\mathbf{y}$ , external prior GMM model	488
<b>Output:</b> The denoised image $\hat{\mathbf{x}}$ .	489
<b>Initialization:</b> $\hat{\mathbf{x}}^{(0)} = \mathbf{y}$ ;	490
<b>for</b> $t = 1 : IteNum$ <b>do</b>	491
1. Extracting internal patch groups (PGs) from $\hat{\mathbf{x}}^{(0)}$ ;	492
<b>for</b> each PG $\mathbf{Y}_n$ <b>do</b>	493
2. Calculate group mean $\mu_n$ and form PG $\bar{\mathbf{Y}}_n$ ;	494
3. Gaussian component selection via Equ. (5);	495
<b>end for</b>	496
<b>for</b> each Internal Subspace <b>do</b>	497
4. Internal Subspace Learning by (4);	498
5. Recover each patch in all PGs via $\hat{\mathbf{x}}_m = \mathbf{D} \hat{\boldsymbol{\alpha}} + \mu_y$ ;	499
<b>end for</b>	500
6. Aggregate the recovered PGs of all subspaces to form the recovered image $\hat{\mathbf{x}}^{(t)}$ ;	501
<b>end for</b>	502

adaptive to the given noisy image. The learning process has closed-form solutions and hence is very efficient. Besides, the learned orthogonal dictionary also makes the denoising process very efficient under sparse coding framework. 3) the singular values obtained by SVD in Equ. (3) reflect the prior weights of the atoms in learned dictionary and can be used as adaptive parameters for real image denoising.

### 3.4. The Denoising Algorithm

As discussed in previous section, the proposed framework employs the external prior learned from natural clean images to guide the internal prior learning for any given noisy image. The image denoising is simultaneously done with the learning process. In the denoising procedure, for each subspace (we ignore the indexes  $k = 1, \dots, K$  of subspaces for notation simplicity), the group mean vectors  $\{\mu_n\}_{n=1}^N$  of the internal noisy PGs  $\{\mathbf{y}_n\}_{n=1}^N$  are saved for reconstruction. Until now, we obtain the solutions of sparse coefficients  $\{\hat{\boldsymbol{\alpha}}_n^{T-1}\}$  in (9) and the orthogonal dictionary  $\mathbf{D}_{(T)} = [\mathbf{D}_e \mathbf{D}_i^{(T)}]$  in Equ. (10). Then the latent clean PGs  $\{\hat{\mathbf{y}}_n\}_{n=1}^N$  are recovered by  $\hat{\mathbf{y}}_n = \mathbf{D} \hat{\boldsymbol{\alpha}}$ . We add back its PG mean denoted as  $\mu_y$  and reconstruct the clean image  $\hat{\mathbf{x}}$  by aggregating all the estimated PGs. Similar to [10], we perform the above denoising procedures for several iterations for better denoising outputs. But different from [10], we do not employ the iterative regularization strategy [35] since it is not effective in our method. The proposed denoising algorithm is summarized in Alg. 1.

## 4. 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

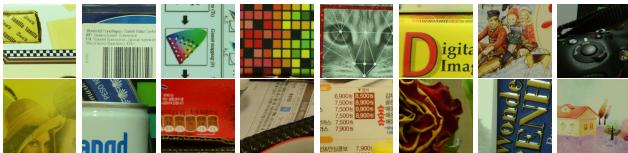


Figure 4. Some cropped images of the dataset [13].

540  
541  
542  
543  
544  
545  
546  
547 by [13], some samples are shown in Figure 5. The sec-  
548 ond dataset is provided by the website of Noise Clinic [19].  
549 The third dataset is provided by the Commercial software  
550 Neat Image [24]. The second and third dataset do not have  
551 ground truth images.  
552

#### 4.1. Implementation Details

553 Our proposed method contains two stages, the external  
554 prior guided internal subspace learning stage and the adap-  
555 tive denoising stage. In the learning stage, there are 4 pa-  
556 rameters: the patch size  $p$ , the number of patches in a PG  
557  $M$ , the window size  $W$  for PG searching and the number of  
558 clusters  $K$ . We set  $p = 6$  (hence the patch size is  $6 \times 6 \times 3$ ),  
559  $M = 10$ ,  $W = 31$ ,  $K = 32$ . We extracted about 3.6  
560 million PGs from the Kodak PhotoCD Dataset, which in-  
561 cludes 24 high quality color images, to train the external  
562 prior via PG-GMM. In the denoising stage, the paramter  
563  $\lambda = 0.002$  is used to regularize the sparse term. The  $\delta$  in  
564 iterative regularization is set as  $\delta = 0.09$ . Given a RGB im-  
565 age of size  $256 \times 256 \times 3$ , we can extract over 60 thousands  
566 local patches of size  $6 \times 6 \times 3$ . It is time-consuming to  
567 search patch groups (PG) for local patches one by one. To  
568 speed up the searching process, we employ the technique of  
569 'Summed Area Table' [34] for efficient PG searching. The  
570 SAT permits to evaluate the sum of pixel values in rectan-  
571 gular regions of the image with four operations, regardless of  
572 the region size. Hence, we do not need do distance measure  
573 for each local patch.  
574

#### 4.2. Comparison on External and Internal methods

575 In this subsection, we compared the proposed external  
576 prior guided internal subspace learning model on real image  
577 denoising. The three methods are evaluated on the dataset  
578 provided in [13]. We calculate the PSNR, SSIM [22] and  
579 visual quality of these three methods. We also compare the  
580 speed. The PSNR and SSIM results on 60 cropped images  
581 from [13] are listed in Table 1. The images are cropped into  
582 size of  $500 \times 500$  for better illustration. We also compare  
583 the three methods on visual quality in Figure 2. Compare  
584 the denoised images listed in Figure 2 and Figure 2, we  
585 can see that the Offline method is better at edges, smooth  
586 regions while the Online method is good at complex tex-  
587 tures. The reason is two folds. Firstly, the Offline method is  
588 learned on clean images and hence is better at representing  
589 edges, structuals, and smooth area. The online method is  
590 influenced by the noise and hence some noise cannot be re-  
591

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

	Noisy	Offline	Online	Guided
PSNR	34.51	38.19	38.07	<b>38.55</b>
SSIM	0.8718	0.9663	0.9625	<b>0.9675</b>

594 moved. Secondly, the Online method is better at recovering  
595 complex area since they could learn adaptive dictionaries  
596 for the specific area. The Offline method cannot recover the  
597 complex area since they did not learn the similar structures  
598 from the external natural clean images.  
599

#### 4.3. Comparison With other Competing Methods

600 We compare with previous state-of-the-art Gaussian  
601 noise removal methods such as BM3D [4], WNNM [8],  
602 MLP [7], CSF [9], and the recently proposed TRD [11].  
603 We also compare with three competing real image denois-  
604 ing methods such as Noise Clinic, Neat Image, and the CC-  
605 Noise method proposed recently. The commercial software  
606 Neat Image [24] first estimates the parameters of noise via  
607 a large flat area and then filters the noise accordingly. All  
608 these methods need noise estimation which is very hard to  
609 perform if there is no uniform regions are available in the  
610 testing image. The NeatImage will fail to perform automati-  
611 cal parameters settings if there is no uniform regions.<sup>1</sup>  
612

613 We the competing denoising methods from various re-  
614 search directions on two datasets. Both the two datasets  
615 comes from the [13]. The first dataset contains 17 images  
616 of size over  $7000 \times 5000$ . Since this dataset contains repet-  
617 itive contents across different images, we crop 60 small im-  
618 ages of size  $500 \times 500$  from these 17 images in [13]. The  
619 PSNR and SSIM resluts are listed in Table 3. The number  
620 in red color and blue color means the best and second best  
621 results, respectively. From the Table 3, we can see that the  
622 external based method can already surpass largely the pre-  
623 vious denoising methods. The improvement on PSNR over  
624 the second best method, i.e., TRD, is 0.44dB. The  
625

#### 4.4. Discussion on Parameter $\lambda$

626 The proposed method only has a key parameter, namely  
627 the regularization paramters  $\lambda$ . To demonstrate that the pro-  
628 posed method is robust to the variance of  $\lambda$ , we vary the  
629 parameter  $\lambda$  across a wide range and obtain the PSNR and  
630 SSIM results as a function of the parameter  $\lambda$ . The re-  
631 sults is shown in Figure 6, from which we can see that the  
632 proposed method can achieve a PSNR (SSIM) over 38.5dB  
633 (0.9660) when  $\lambda$  varies from 0.0015 to 0.0025. This shows  
634 that the proposed method is indeed robust to the chosen of  
635 the paramter  $\lambda$ .  
636

<sup>1</sup>To compare with CCNoise, we first transform the denoised images into double format.  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647

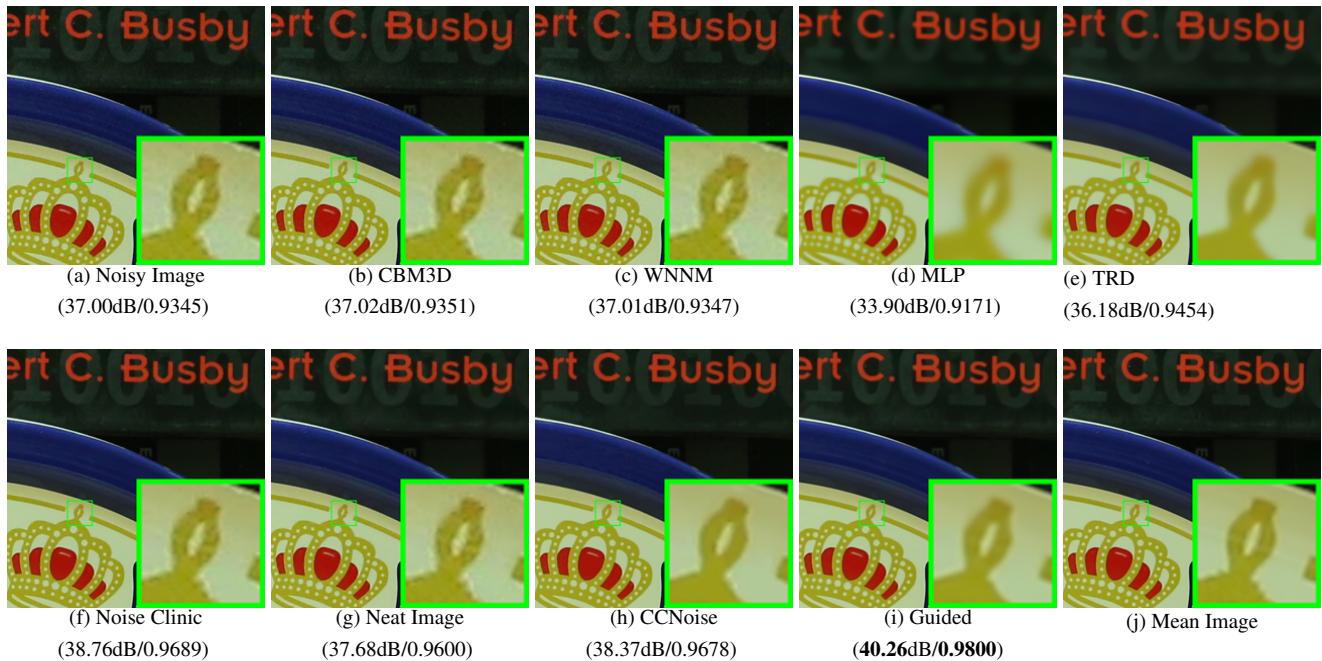


Figure 5. Denoised images of the image "Canon 5D Mark 3 ISO 3200 1" by different methods. The images are better to be zoomed in on screen.

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

	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

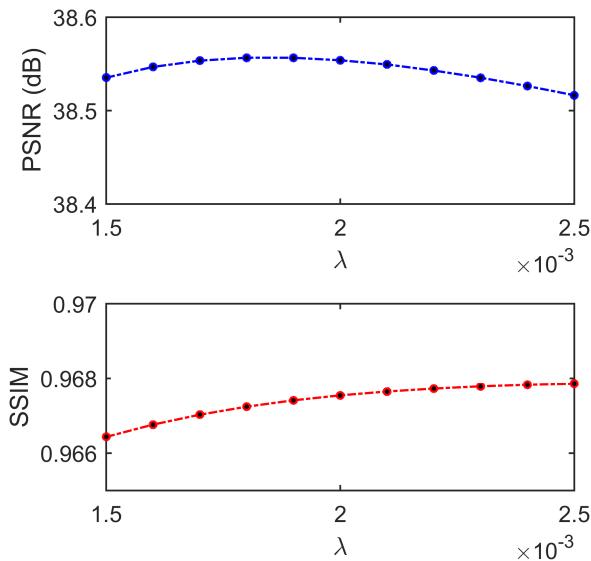


Figure 6. The PSNR/SSIM results as a function of the parameter  $\lambda$ .

## 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
- [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, 5
- [4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collabora-

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

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

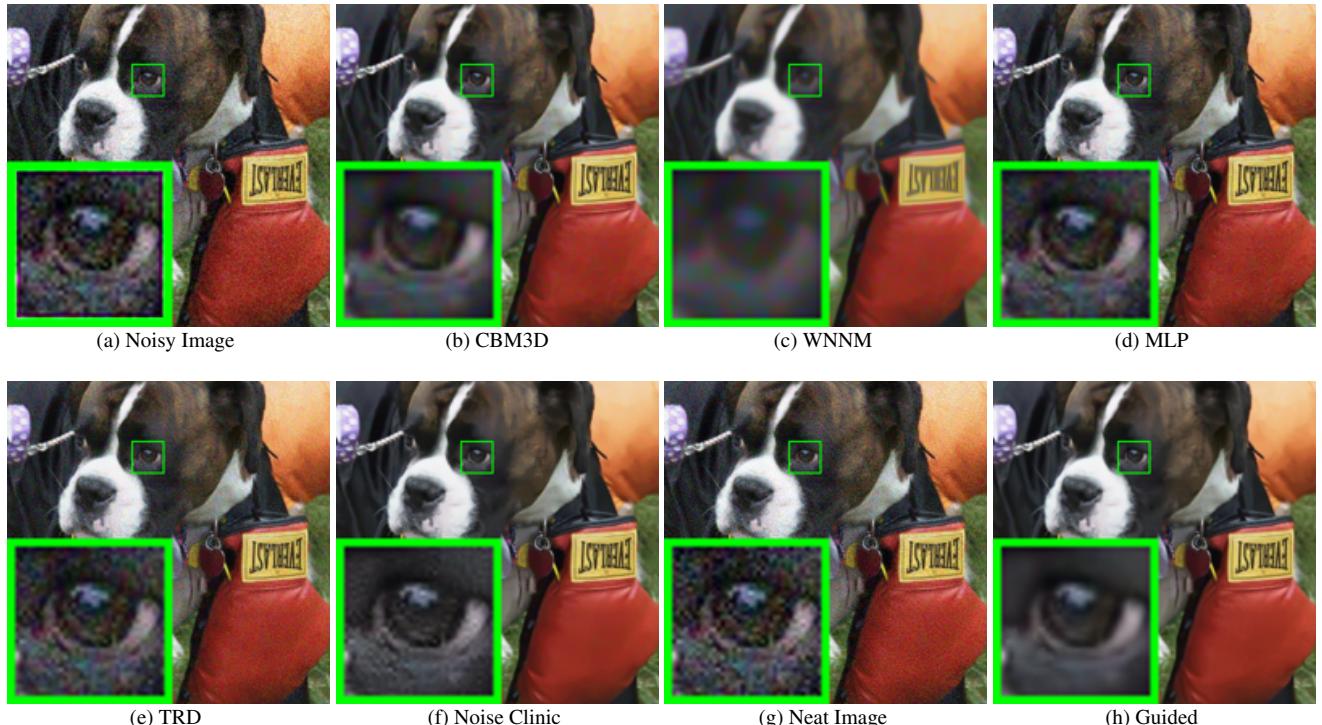


Figure 7. Denoised images of the image "5dmak3iso32003" by different methods. The images are better to be zoomed in on screen.

tive 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

- 864 [8] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nu- 918  
865 clear norm minimization with application to image denois- 919  
866 ing. *CVPR*, pages 2862–2869, 2014. 1, 2, 3, 6 920  
867  
868 [9] U. Schmidt and S. Roth. Shrinkage fields for effective 921  
image restoration. *Computer Vision and Pattern Recognition 922* (CVPR), 2014 IEEE Conference on, pages 2774–2781, June 923  
2014. 1, 2, 6 924  
869  
870 [10] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng. Patch 925  
group based nonlocal self-similarity prior learning for image 926  
denoising. *2015 IEEE International Conference on Computer 927  
Vision (ICCV)*, pages 244–252, 2015. 1, 2, 3, 4, 5 928  
871  
872 [11] Yunjin Chen, Wei Yu, and Thomas Pock. On learning 929  
optimized reaction diffusion processes for effective image 930  
restoration. *Proceedings of the IEEE Conference on Computer 931  
Vision and Pattern Recognition*, pages 5261–5269, 2015. 1, 2, 6 932  
873  
874  
875  
876 [12] S. J. Kim, H. T. Lin, Z. Lu, S. Ssstrunk, S. Lin, and M. S. 933  
Brown. A new in-camera imaging model for color computer 934  
vision and its application. *IEEE Transactions on Pattern 935  
Analysis and Machine Intelligence*, 34(12):2289–2302, Dec 936  
2012. 1 937  
877  
878 [13] Seonghyeon Nam, Youngbae Hwang, Yasuyuki Matsushita, 938  
and Seon Joo Kim. A holistic approach to cross-channel im- 939  
age noise modeling and its application to image denoising. 940  
*Proc. Computer Vision and Pattern Recognition (CVPR)*, 941  
pages 1683–1691, 2016. 1, 2, 3, 6, 7, 8 942  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917 [14] Glenn E Healey and Raghava Kondepudy. Radiometric ccd 943  
camera calibration and noise estimation. *IEEE Transactions 944  
on Pattern Analysis and Machine Intelligence*, 16(3):267– 945  
276, 1994. 1 946  
[15] J. Portilla. Full blind denoising through noise covariance 947  
estimation using gaussian scale mixtures in the wavelet do- 948  
main. *Image Processing, 2004. ICIP '04. 2004 International 949  
Conference on*, 2:1217–1220, 2004. 1, 3 950  
[16] Tamer Rabie. Robust estimation approach for blind denois- 951  
ing. *Image Processing, IEEE Transactions on*, 14(11):1755– 952  
1765, 2005. 1, 3 953  
[17] C. Liu, R. Szeliski, S. Bing Kang, C. L. Zitnick, and W. T. 954  
Freeman. Automatic estimation and removal of noise from 955  
a single image. *IEEE Transactions on Pattern Analysis and 956  
Machine Intelligence*, 30(2):299–314, 2008. 1, 3 957  
[18] Zheng Gong, Zuowei Shen, and Kim-Chuan Toh. Image 958  
restoration with mixed or unknown noises. *Multiscale Mod- 959  
eling & Simulation*, 12(2):458–487, 2014. 1, 3 960  
[19] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image 961  
blind denoising. *Image Processing, IEEE Transactions on*, 962  
24(10):3149–3161, 2015. 1, 2, 3, 6 963  
[20] Fengyuan Zhu, Guangyong Chen, and Pheng-Ann Heng. 964  
From noise modeling to blind image denoising. *The IEEE 965  
Conference on Computer Vision and Pattern Recognition 966  
(CVPR)*, June 2016. 1, 3 967  
[21] C. M. Bishop. *Pattern recognition and machine learning*. 968  
New York: Springer, 2006. 1 969  
[22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. 970  
Image quality assessment: from error visibility to struc- 971  
tural similarity. *IEEE Transactions on Image Processing*, 972  
13(4):600–612, 2004. 1, 6 973  
[23] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Color 974  
image denoising via sparse 3d collaborative filtering with 975  
grouping constraint in luminance-chrominance space. *IEEE 976  
International Conference on Image Processing*, 1, 2007. 2 977  
[24] Neatlab ABSsoft. Neat image. [https://ni.  
neatvideo.com/home](https://ni.neatvideo.com/home). 2, 6 978  
[25] G. Yu, G. Sapiro, and S. Mallat. Solving inverse problems 979  
with piecewise linear estimators: From Gaussian mixture 980  
models to structured sparsity. *IEEE Transactions on Image 981  
Processing*, 21(5):2481–2499, 2012. 2, 3 982  
[26] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally central- 983  
ized sparse representation for image restoration. *Image Pro- 984  
cessing, IEEE Transactions on*, 22(4):1620–1630, 2013. 2, 3 985  
[27] J. Portilla, V. Strela, M.J. Wainwright, and E.P. Simoncelli. 986  
Image denoising using scale mixtures of Gaussians in the 987  
wavelet domain. *Image Processing, IEEE Transactions on*, 988  
12(11):1338–1351, 2003. 3 989  
[28] Peter J Huber. *Robust statistics*. Springer, 2011. 3 990  
[29] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian 991  
image denoising algorithm. *SIAM Journal on Imaging Sci- 992  
ences*, 6(3):1665–1688, 2013. 3 993  
[30] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum 994  
likelihood from incomplete data via the EM algorithm. *Journal 995  
of the Royal Statistical Society. Series B (methodological)*, 996  
pages 1–38, 1977. 4, 5 997  
[31] David L Donoho and Michael Elad. Optimally sparse repre- 998  
sentation in general (nonorthogonal) dictionaries via 1 mini- 999  
mization. *Proceedings of the National Academy of Sciences*, 100(5):2197–2202, 2003. 1000  
[32] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. 1001  
Inf. Theor.*, 41(3):613–627, 1995. 1002  
[33] Chenglong Bao, Jian-Feng Cai, and Hui Ji. Fast sparsity- 1003  
based orthogonal dictionary learning for image restoration. 1004  
*Proceedings of the IEEE International Conference on Com- 1005  
puter Vision*, pages 3384–3391, 2013. 5 1006  
[34] Franklin C. Crow. Summed-area tables for texture mapping. 1007  
*SIGGRAPH Comput. Graph.*, 18(3):207–212, January 1984. 1008  
6 1009  
[35] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An it- 1010  
erative regularization method for total variation-based image 1011  
restoration. *Multiscale Modeling & Simulation*, 4(2):460– 1012  
489, 2005. 5 1013  
[36] C. M. Bishop. *Pattern recognition and machine learning*. 1014  
New York: Springer, 2006. 1 1015