

000
001
002054
055
056

External Patch Group Prior Guided Internal Subspace Learning for Real Image Denoising

057
058
059008
009
010
011
012060
061
062
063

Anonymous CVPR submission

013
014
015064
065
066

Paper ID ****

067
068
069016
017
018
019
020070
071
072
073
074

Abstract

Existing image denoising methods largely depends on noise modeling and estimation. The commonly used noise models, additive white Gaussian, are inflexible in describing the complex noise on real noisy images. This would limit the performance of existing methods on denoising real noisy images. In this paper, we firstly demonstrate that almost all state-of-the-art methods on removing Gaussian noise and real noise are limited in denoising real noisy images. We demonstrate that a simple Patch Group based Prior Learning model on RGB images can achieve better performance than existing denoising methods, especially the ones designed for real noise in natural images. Besides, we employ the external patch group prior learning for internal clustering and subspace learning. This external information guided internal denoising methods achieves even better than the external PG prior based methods and the fully internal PG prior based method. Through extensive on standard datasets on real noisy images with groundtruth, we demonstrate that the proposed method achieves much better denoising performance than the other state-of-the-art methods on Gaussian noise removal and real noise removal.

037
038
039075
076
077

1. Introduction

040
041
042
043
044078
079
080
081
082

Image denoising is a fundamental problem in computer vision and image processing. It is 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). However, in real world, the noise in real images are much more complex than Gaussian. The cameras will undertake high ISO settings for high-speed shots on actions, long exposure for low light on night shots, etc. These conditions would produce noise as a by-product. Besides,

the in-camera imaging pipeline [12, 13] would also change the noise distribution. Therefore, the noise in real images are much more complex than Gaussian, and depends on camera series, brands, as well as the settings (ISO, shutter speed, and aperture, etc). The models designed to deal with AWGN would become much less effective on real noisy images.

In the last decade, the methods of [14, 15, 16, 17, 18, 19, 13] are designed to deal with real noisy images. Almost all these methods coincidently employ a two-stage framework: in the first stage, assuming a distribution model (usually Gaussian) on the noise and estimate its parameters; in the second stage, 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 [16]. Although the mixture of Gaussians (MoG) model is possible to approximate any unknown noise [19], estimating its parameters is often time consuming via nonparametric Bayesian techniques [19, 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]. This image is captured by a Nikon D800 camera while the ISO is set as 3200. The "ground truth" image is also provided with which we can calculate objective measurements. More details about this dataset can be found at the experimental section. 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. This proves that the above mentioned methods are not effective on denoising complex noise on real images.

In this paper, we attempt to deal with complex noise in real images by integrating the external and internal information. Since the real noise is signal dependent [13, 23], the prior information in external natural images can be employed to avoid the high correlation between noise and signal in internal images. On the other hand, the internal prior is adaptive to the image and can recover better the latent clean image. Based on these observations, we

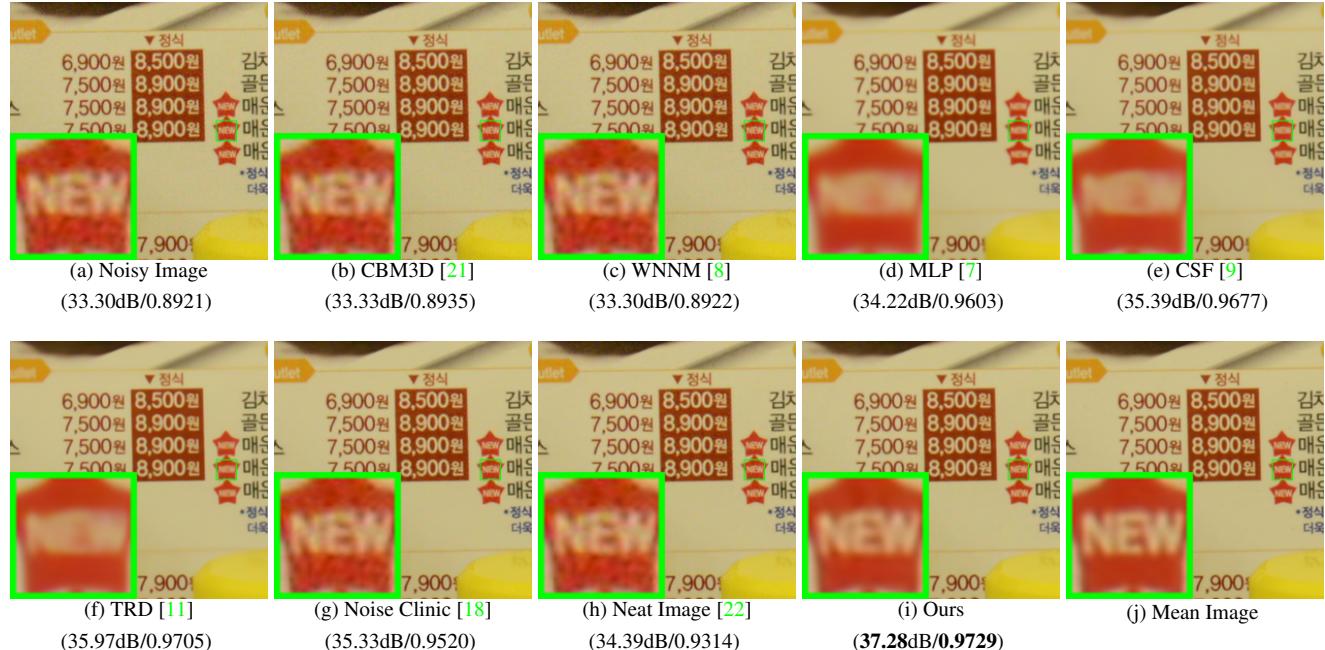


Figure 1. Denoised images of the real noisy image “*NikonD800ISO3200A3*” by different methods. The images are better to be zoomed on screen.

make detailed study on internal and external information for real image denoising task. We made several observations. Firstly, we found that the Patch Group Prior learning based denoising method [10] learned on clean RGB images are enough to outperform the above mentioned denoising methods. Secondly, we also found that a fully internal PG prior based denoising method which achieve better performance than the fully external method. Most importantly, we found that the external PG prior guided internal method can achieve even better and faster performance on real image denoising. In fact, the external PG prior learning based model is employed to guide the clustering of internal PGs extracted from the input noisy images. Then for each cluster of PGs, we perform subspace learning by PCA and denoising by weighted sparse coding. We perform comprehensive experiments on real noisy images captured by different CMOS or CCS sensors. The results demonstrate that our method achieves comparable or even better performance on denoising real noisy images. An initial glimpse of our method is also listed in Figure 1. This reveals the potential advantages of combining external and internal information of natural images on robust and complex real noisy image denoising problem.

1.1. Our Contributions

The contributions of this paper are summarized as follows:

- We propose a novel method which combine the external and internal PG prior for real noisy image denois-

ing problem;

- Our method doesn’t need noise modeling and estimation, and the noise levels of real noisy images are automatically expressed by the singular values of learned subspace;
- We achieve much better performance on visual quality, PSNR, SSIM, and speed, than other competing methods for real image denoising problem.

The rest of this paper will be summarized as follows: in Section 2, we will introduce the related work close to our work; in Section 3, we will introduce our proposed external prior guided internal subspace learning framework for real image denoising; in Section 4, we will demonstrate the denoising experiments on several standard dataset; in Section 5, we will conclude our paper and give our 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]. The better modeling on NSS is demonstrated via better image denoising performance on natural images. However, in [10], only the PGs of clean natural

216 images is utilized, while the PGs of noisy input images are
 217 ignored. In this paper, we aim at making use of both PGs
 218 from natural clean images and real noisy images for better
 219 denoising performance.
 220

221 2.2. Internal v.s. External Image Denoising

223 For natural images, the internal patch recurrence across
 224 multiple scales has been successfully applied in many
 225 image restoration problems [24, 25, 26, 27]. These
 226 work demonstrate that internal information is enough for
 227 many ill-posed problems including denoising additive white
 228 Gaussian noise. The rationale is, since the AWGN noise is
 229 independent of the original clean images, it will be reduced
 230 if the image is scaled to a smaller size. However, the noise
 231 in real images is generated mostly from the camera sensors,
 232 which is highly complex and signal dependent [13]. Be-
 233 sides, according to the seminar workd of [23], the noise in
 234 real images has fixed patterns from several main sources.
 235 Therefore, we can hardly separate the complex noise from
 236 the signals without the help of external (correct) informa-
 237 tion of natural clean images. Only using the internal infor-
 238 mation may be not enough for real image denoising prob-
 239 lem. On the other hand, the methods only using the exter-
 240 nal information may be not adaptive for real noisy images.
 241 Recently, the methods of [7, 9, 11] had been proposed to
 242 learn, on both internal and external images, a process di-
 243 rectly mapping the noisy patches to denoised ones. These
 244 discriminative learning based methods are only effective on
 245 additive white Gaussian noise and ineffective on complex
 246 and signal dependent noise in real images. This has been
 247 shown in Figure 1. In this paper, our goal is to make use
 248 of external information to guide the subspace learning of
 249 internal PGs for image denoising task.

250 2.3. Real Image Denoising

251 To the best of our knowledge, the study of real image
 252 denoising can be dated back to the BLS-GSM model [28],
 253 in which Portilla et al. proposed to use scale mixture of
 254 Gaussian in overcomplete oriented pyramids to estimate the
 255 latent clean images. In [14], Portilla proposed to use a cor-
 256 related Gaussian model for noise estimation of each wavelet
 257 subband. Based on the robust statistics theory [?], the work
 258 of Rabie [15] modeled the noisy pixels as outliers, which
 259 could be removed via Lorentzian robust estimator. In [16],
 260 Liu et al. proposed to use ‘noise level function’ (NLF) to
 261 estimate the noise and then use Gaussian conditional ran-
 262 dom field to obtain the latent clean image. Recently, Gong
 263 et al. proposed an optimization based method [17], which
 264 models the data fitting term by weighted sum of ℓ_1 and ℓ_2
 265 norms and the regularization term by sparsity prior in the
 266 wavelet transform domain. Later, Lebrun el al. proposed
 267 a multiscale denoising algorithm called ‘Noise Clinic’ [18]
 268 for real image denoising task. This method generalizes the
 269

270 NL-Bayes [29] to deal with signal, scale, and frequency de-
 271 pendent noise. Recently, Zhu et al. proposed a Bayesian
 272 model [19] which approximates the noise via Mixture of
 273 Gaussian (MoG) model [20]. The clean image is recovered
 274 from the noisy image by the proposed Low Rank MoG fil-
 275 ter (LR-MoG). In this paper, we proposed a noval denoising
 276 method achieving much better performance than previous
 277 real denoising methods.
 278

279 3. External Patch Group Prior Guided Inter- 280 nal Subspace Learning

281 In this section, we formulate the framework of external
 282 Patch Group prior guided internal subspace learning. We
 283 first introduce the patch group prior leaning on clean natural
 284 RGB images. Then we formulate the external guided inter-
 285 nal subspace learning. Finally, we discuss the differences
 286 between external subspaces and the corresponding internal
 287 subspace.
 288

289 3.1. External Patch Group Prior Learning

290 Natural images often demonstrate repetitive patterns,
 291 this nonlocal self-similarity (NSS) property is a key suc-
 292 cessful factor for many image denoising methods [1, 4, 5,
 293 30, 8, 10] and restoration methods []. In [10], the NSS prop-
 294 erty is directly learned as an external prior in a patch group
 295 manner. In this section, we formulate the Patch Group prior
 296 on natural color images.
 297

298 In [10], the patch group (PG) is defined as a group of
 299 similar patches to the local patch. The patch group mean
 300 is destroyed, and hence different groups patches can share
 301 similar PGs. Therefore the space to be modeled is largely
 302 reduced. In this work, we extract PGs from RGB im-
 303 ages. Each patch is of size $p \times p \times 3$. For each local
 304 patch, we search its similar patches around it through the
 305 Euclidean distance in a local window of size $W \times W$. The
 306 PG is denoted by $\{\mathbf{x}_m\}_{m=1}^M$, where $\mathbf{x}_m \in \mathbb{R}^{3p^2 \times 1}$
 307 is a color image patch vector. The mean vector of this
 308 PG is $\mu = \frac{1}{M} \sum_{m=1}^M \mathbf{x}_m$, and $\bar{\mathbf{x}}_m = \mathbf{x}_m - \mu$ is the
 309 group mean subtracted patch vector. The PG is defined as
 310 $\bar{\mathbf{X}} \triangleq \{\bar{\mathbf{x}}_m\}, m = 1, \dots, M$, and it represent the external
 311 NSS prior on color images. Assume we have extracted N
 312 PGs from a given set of natural images, and the n -th PG
 313 is defined as $\bar{\mathbf{X}}_n \triangleq \{\bar{\mathbf{x}}_{n,m}\}_{m=1}^M, n = 1, \dots, N$. We employ
 314 the patch group based Gaussian Mixture Model (PG-GMM)
 315 for NSS prior learning. We aim to learn a set of K Gaus-
 316 sians $\{\mathcal{N}(\mu_k, \Sigma_k)\}$ from N training PGs $\{\bar{\mathbf{X}}_n\}$, while re-
 317 quiring that all the M patches $\{\bar{\mathbf{x}}_{n,m}\}$ in PG $\bar{\mathbf{X}}_n$ belong to
 318 the same Gaussian component and assume that the patches
 319 in the PG are independently sampled. Note that such an as-
 320 sumption is commonly used in patch based image modeling
 321 [3, 5]. Then, the likelihood of $\{\bar{\mathbf{X}}_n\}$ can be calculated as
 322

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

324 By assuming that all the PGs are independently sampled,
 325 the overall objective log-likelihood function is
 326

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

327 We maximize the above objective function for PG-GMM
 328 learning and finally obtain the GMM model with learned pa-
 329 rameters including mixture weights $\{\pi_k\}_{k=1}^K$, mean vectors
 330 $\{\boldsymbol{\mu}_k = \mathbf{0}\}_{k=1}^K$, and covariance matrices $\{\boldsymbol{\Sigma}_k\}_{k=1}^K$. Noted
 331 that the mean vector of each cluster is natural zeros, i.e.,
 332 $\boldsymbol{\mu}_k = \mathbf{0}$.
 333

334 3.2. External Prior Guided Internal Subspace 335 Learning

336 Given a rael noisy image, we extract noisy PGs from it
 337 and save the mean vectors of each PG for recovering. The
 338 mean substracted PG is defined as $\bar{\mathbf{Y}}$. To project this PG
 339 into a most adaptive subspace, we select the most suitable
 340 Gaussian component to it from the PG-GMM trained in pre-
 341 vious section. The selection can be done by checking the
 342 posterior probability that $\bar{\mathbf{Y}}$ belongs to the k th Gaussian
 343 component:

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

344 Since the noise on real images are mostly small when com-
 345 pared to the signals, the covariance matrix of the k th com-
 346 ponent is still $\boldsymbol{\Sigma}_k$. Finally, the component with the maxi-
 347 mum A-posteriori (MAP) probability $\ln P(k|\bar{\mathbf{Y}})$ is selected
 348 as the most suitable subspace for $\bar{\mathbf{Y}}$.

349 Though each PG has been projected into its most suitable
 350 subspace, the pre-learned subspace is still too general to
 351 represent the noisy PG extracted from the real noisy image.
 352 That is, the noisy PGs projected into one cluster can still
 353 constisted a subspace which is of lower dimensions than the
 354 subspace pre-learned from the external PGs. This can be
 355 demonstrated by compare the distribution of external PGs
 356 and internal PGs in the same clusters. We randomly select
 357 one cluster, and collect the celan PGs extracted from exter-
 358 nal dataset (Kodak 24 images) and the niosy PGs from the
 359 testing image. Since the original PGs are of $3p^2$ dimensions,
 360 we apply PCA to project the PGs into 2 dimensions for bet-
 361 ter visualization. The results is shown in Figure ??, from
 362 which we can see clearly that the projected PGs are mainly
 363 in a smaller region of the external PGs, which proves that
 364 the internal PGs are only consisted a subspace in a lower
 365 dimension than the PGs collected from external subspace.
 366 To better and adaptively charactering the internal PGs from
 367 the testing image, we need learn a more specific dictionary
 368 for noisy PGs assigned into each cluster. For notation sim-
 369 plicity, we ignore the index of subspace k . The internal PGs
 370

371 \mathbf{Y} form a subspace which can be obtained by singular value
 372 decomposition (SVD),
 373

$$\mathbf{Y} = \mathbf{USV}^T \quad (4)$$

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

378 3.3. What does External Prior help the Internal 379 Subspace Learning?

380 Until now, we have divided the noisy PGs into multiple
 381 internal subspaces. Here we take a deep analysis on how the
 382 external NSS prior guide the subspace learning of internal
 383 PGs. The help are at least threefold. Firstly, through MAP
 384 in (3), the external prior guides the noisy PGs to be clustered
 385 into the correct subspaces. If we cluster the noisy PGs in an
 386 automatical way, the subspaces we learned will be highly
 387 degraded by the signal dependent noise. Secondly, the guid-
 388 ance of external prior for internal clustering is more efficient
 389 than directly clustering the internal noisy PGs. It only needs
 390 to calculate the MAP probability via the equation (3) while
 391 the internal clustering via GMM is time-consuming on EM
 392 algorithm [31]. Thirdly, due to the correct guidance of ex-
 393 ternal prior, the strucual decompositon via SVD of each
 394 subspace is more adaptive. This will bring better denoising
 395 performance than the methods only using the external
 396 information.

397 3.4. Discussion on Internal Subspace Structure

398 The Internal PGs are in fact lying in the subspaces of
 399 external PG Spaces. To defend this argument, we com-
 400 pare the distribution of external PGs extracted from clean
 401 natural images and real noisy images. For better illumina-
 402 tion, we randomly selected a cluster and project the origi-
 403 nal clean PGs \mathbf{X} onto a 2-D plane. This could be done
 404 via $\mathbf{X}_p = \mathbf{U}(:, 1 : 2)^T \mathbf{X}$, where \mathbf{U} is the singular vector
 405 matrix of that cluster. The noisy PGs \mathbf{Y} assigned in this
 406 cluster is also projected into 2-D via $\mathbf{Y}_p = \mathbf{U}(:, 1 : 2)^T \mathbf{Y}$.
 407 The Figure ?? reflects the distribution on the 2-D plane of
 408 the projected clean PGs from external natural images and
 409 the projected noisy PGs from internal image. We can see
 410 that the internal noisy PGs are indeed lying in a subspace
 411 of the external PGs. Hence, if we directly use the external
 412 prior learned from clean PGs, the learned subspaces would
 413 be too generative to be suitable for the testing data.

414 Through SVD, the PGs in each internal subspace can be
 415 divided into singular vectors and singular values. The sin-
 416 gular vectors are the basis of the corresponding subspace
 417 while the singular values reflect the importance of these ba-
 418 sis. The basis can be used as dictionary to code the noisy
 419 PGs. And the sigular values are adaptive parameters for in-
 420 ternal noisy PGs. We can compare the singular values of

432 one internal subspace and the corresponding space of external PGs. The result is shown in Figure ???. From which
 433 we can see that the noisy subspace often have higher values than external space consisted of clean PGs. This gap is
 434 clearly made of the noise and can be used for image denoising
 435 in a natural way.
 436
 437

4. The Denoising Algorithm

4.1. Adaptive Sparse Coding

We employ the \mathbf{U} as the adaptive dictionary to represent the structural variations of the PGs in that component.

$$\min_{\alpha} \|\bar{\mathbf{y}}_m - \mathbf{U}\alpha\|_2^2 + \sum_{i=1}^{3p^2} \lambda_i |\alpha_i|. \quad (5)$$

The i th entry of the regularization parameter λ_i

$$\lambda_i = \lambda / (\mathbf{S}_i + \varepsilon), \quad (6)$$

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 (4) has a closed-form solution (detailed derivation can be found in the supplementary material):

$$\hat{\alpha} = \text{sgn}(\mathbf{U}^T \bar{\mathbf{y}}_m) \odot \max(|\mathbf{U}^T \bar{\mathbf{y}}_m| - \Lambda, 0), \quad (7)$$

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.

4.2. Fast Patch Group Searching by Integral Image

The searching of patch groups in images is inefficient if we search non-local similar patches to each local patch. To speed up the searching process and make our proposed method faster, we employ the technique of 'Summed Area Table' [32] 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[33]

4.3. Effectively dealing with different noisy images

For real image denoising, we can perform well on images which have similar noise levels with the training dataset. How can we deal with the real noisy images whose noise levels are higher than the training dataset? The answer is to remove the noise by more iterations. The input image of each iteration is the recovered image of previous iteration. This makes sense since we can still view the recovered image as a real noisy image.

Alg. 1: External Guided Internal Subspace Denoising	486
Input: Noisy image \mathbf{y} , PG-GMM model	487
1. Initialization: $\hat{\mathbf{x}}^{(0)} = \mathbf{y}, \mathbf{y}^{(0)} = \mathbf{y};$	488
for $t = 1 : IteNum$ do	489
2. Iterative Regularization:	490
$\mathbf{y}^{(t)} = \hat{\mathbf{x}}^{(t-1)} + \delta(\mathbf{y} - \mathbf{y}^{(t-1)});$	491
3. Estimate the standard deviation of noise;	492
for each PG \mathbf{Y} do	493
4. Calculate group mean μ_y and form PG $\bar{\mathbf{Y}}$;	494
5. Gaussian component selection via (3);	495
end for	496
for each Internal Subspace do	497
6. Internal Subspace Learning by (4);	498
7. Recover each patch in all PGs via $\hat{\mathbf{x}}_m = \mathbf{D}\hat{\alpha} + \mu_y$;	499
end for	500
8. Aggregate the recovered PGs of all subspaces to form	501
the recovered image $\hat{\mathbf{x}}^{(t)}$;	502
end for	503
Output: The recovered image $\hat{\mathbf{x}}^{(IteNum)}$.	504

This will also bring a second problem, that how we could automatically terminate the iteration. This can be solved by two methods. One way is to compare the images between two iterations and calculate their difference, the iteration can be terminated if the difference is smaller than a threshold. The other way is to estimate the noise level of the current image and terminate the iterations when the noise level is lower than a preset threshold. We employ the second way and set the threshold as 0.002 in our experiments. In fact, most of our testing images will be denoised well in one iteration.

4.4. The Overall Algorithm

With the solution $\hat{\alpha}$ in (7), 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 [34] 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 [18]. The third dataset is provided by the Commercial software Neat Image [22]. The second and third dataset do not have ground truth images.



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



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

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 parameter $\lambda = 0.002$ is used to regularize the sparse term. The δ in iterative regularization is set as $\delta = 0.09$.

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 [?] 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, structures, 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

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	38.55
SSIM	0.8718	0.9663	0.9625	0.9675

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 CCNoise method proposed recently. The popular software NeatImage which is one of the best denoising software available. All these methods need noise estimation which is very hard to perform if there is no uniform regions 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 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 .

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



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.

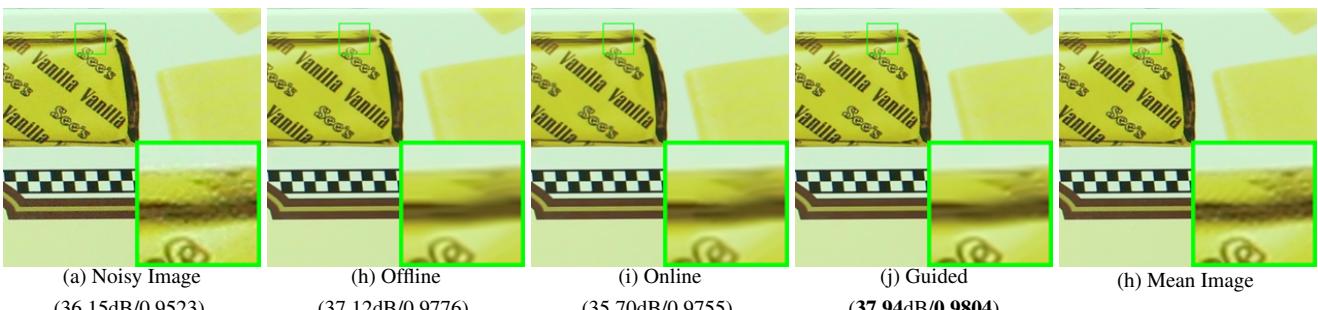


Figure 5. Denoised images of the image "Canon EOS 5D Mark3 ISO 3200 C1" by different methods. The images are better to be zoomed in on screen.

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](#), [3](#)
- [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](#), [3](#), [6](#)
- [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](#), [3](#), [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](#), [3](#), [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*

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 Analysis and Machine Intelligence, 34(12):2289–2302, Dec
785 2012. 1

- 786 [13] Seonghyeon Nam, Youngbae Hwang, Yasuyuki Matsushita,
787 and Seon Joo Kim. A holistic approach to cross-channel im-
788 age noise modeling and its application to image denoising.
789 *Proc. Computer Vision and Pattern Recognition (CVPR)*,
790 pages 1683–1691, 2016. 1, 3, 5, 6, 8
- 791 [14] J. Portilla. Full blind denoising through noise covariance
792 estimation using gaussian scale mixtures in the wavelet do-
793 main. *Image Processing, 2004. ICIP '04. 2004 International
794 Conference on*, 2:1217–1220, 2004. 1, 3
- 795 [15] Tamer Rabie. Robust estimation approach for blind denoising.
796 *Image Processing, IEEE Transactions on*, 14(11):1755–
797 1765, 2005. 1, 3
- 798 [16] C. Liu, R. Szeliski, S. Bing Kang, C. L. Zitnick, and W. T.
799 Freeman. Automatic estimation and removal of noise from
800 a single image. *IEEE Transactions on Pattern Analysis and
801 Machine Intelligence*, 30(2):299–314, 2008. 1, 3
- 802 [17] Zheng Gong, Zuowei Shen, and Kim-Chuan Toh. Image
803 restoration with mixed or unknown noises. *Multiscale Mod-
804eling & Simulation*, 12(2):458–487, 2014. 1, 3
- 805 [18] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image
806 blind denoising. *Image Processing, IEEE Transactions on*,
807 24(10):3149–3161, 2015. 1, 2, 3, 5
- 808 [19] Fengyuan Zhu, Guangyong Chen, and Pheng-Ann Heng.
809 From noise modeling to blind image denoising. *The IEEE*

838 Conference on Computer Vision and Pattern Recognition
839 (CVPR), June 2016. 1, 3

- 840 [20] C. M. Bishop. *Pattern recognition and machine learning*.
841 New York: Springer, 2006. 1, 3
- 842 [21] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Color
843 image denoising via sparse 3d collaborative filtering with
844 grouping constraint in luminance-chrominance space. *IEEE
845 International Conference on Image Processing*, 1, 2007. 2
- 846 [22] Neatlab ABSsoft. Neat image. <https://neatvideo.com/home>. 2, 5
- 847 [23] Glenn E Healey and Raghava Kondepudy. Radiometric ccd
848 camera calibration and noise estimation. *IEEE Transactions
849 on Pattern Analysis and Machine Intelligence*, 16(3):267–
850 276, 1994. 1, 3
- 851 [24] Daniel Glasner, Shai Bagon, and Michal Irani. Super-
852 resolution from a single image. *ICCV*, 2009. 3
- 853 [25] Maria Zontak, Inbar Mosseli, and Michal Irani. Separating
854 signal from noise using patch recurrence across scales. *Pro-
855 ceedings of the IEEE Conference on Computer Vision and
856 Pattern Recognition*, pages 1195–1202, 2013. 3
- 857 [26] Tomer Michaeli and Michal Irani. Blind deblurring using in-
858 ternal patch recurrence. *European Conference on Computer
859 Vision*, pages 783–798, 2014. 3
- 860 [27] Y. Bahat and M. Irani. Blind dehazing using internal patch
861 recurrence. *IEEE International Conference on Computa-
862 tional Photography (ICCP)*, pages 1–9, May 2016. 3

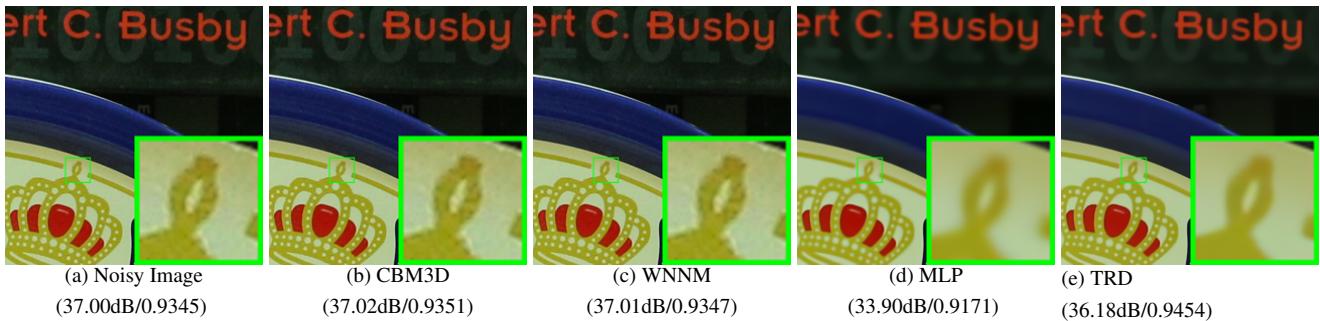


Figure 6. Denoised images of the image "Canon 5D Mark 3 ISO 3200 1" by different methods. The images are better to be zoomed in on screen.

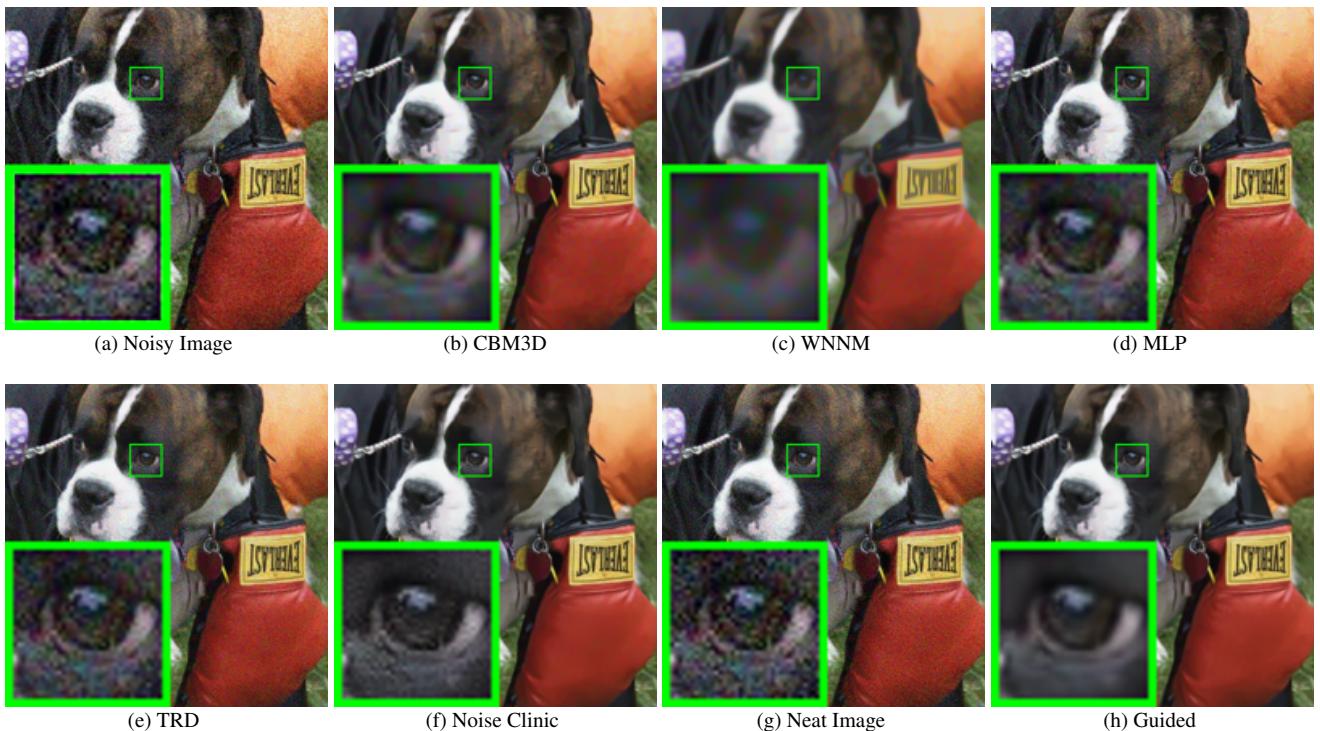
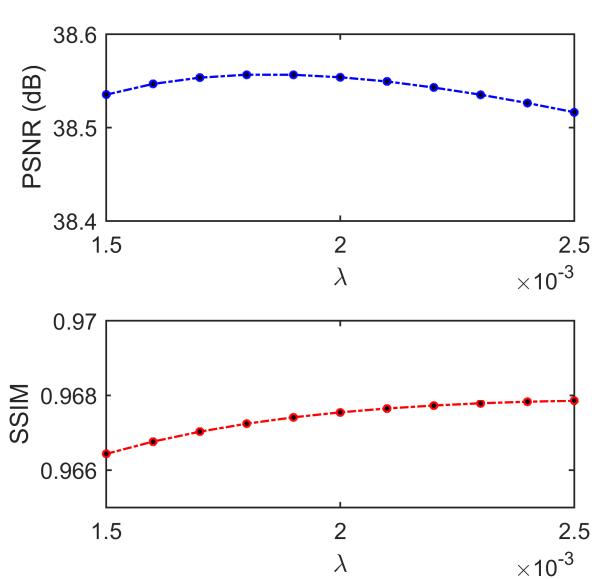


Figure 7. Denoised images of the image "5dmak3iso32003" by different methods. The images are better to be zoomed in on screen.

[28] 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*, 12(11):1338–1351, 2003. 3

[29] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3):1665–1688, 2013. 3

[30] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally central-



991 Figure 8. The PSNR/SSIM results as a function of the parameter
992 λ .
993

994
995 ized sparse representation for image restoration. *Image Pro-*
996 *cessing, IEEE Transactions on*, 22(4):1620–1630, 2013. 3

997 [31] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum
998 likelihood from incomplete data via the EM algorithm. *Journal*
999 *of the Royal Statistical Society. Series B (methodologi-*
1000 *cal)*, pages 1–38, 1977. 4

1001 [32] Franklin C. Crow. Summed-area tables for texture mapping.
1002 *SIGGRAPH Comput. Graph.*, 18(3):207–212, January 1984.
1003 5

1004 [33] Paul Viola and Michael J Jones. Robust real-time face detec-
1005 tion. *International journal of computer vision*, 57(2):137–
1006 154, 2004. 5

1007 [34] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An it-
1008 erative regularization method for total variation-based image
1009 restoration. *Multiscale Modeling & Simulation*, 4(2):460–
1010 489, 2005. 5

1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025

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