A Noise Estimation Free Framework for Robust Real Image Denoising

Anonymous CVPR submission

Paper ID ****

Abstract

Noise modeling and estimation have been two key factors for the success of tranditional image denoising methods. The noise is often assumed to be homogeneous white Gaussian distributed. However, noise in real images are much more complex. Though the Mixture of Gaussians model is employed for modeling more complex noises, the estimation process is unavoidedly time consumming. In this paper, we propose a noval method for real image denoising without noise modeling or estimation. Different from previous methods, we formulate the real image denoising as a cross-domain image synthesis problem. We also propose a bi-coupled dictioanry learning model, in which the coefficients over dictionaries on both domains are coupled by two important regularities relationships. Experiments show that the proposed Bi-Coupled Dictionary Learing (BCDL) model is very effctive in real image denoising problem, resulting competing performance on quality measurements such as visual quality, PSNR, and SSIM.

1. Introduction

Image denoising is a fundermental problem in computer vision and image processing. This problem has been studied for several decades. It is an ideal platform for testing natural image models and an important part of image enhancement for digital image acquisition systems. It also provides high-quality images and better features for other conputer vision tasks such as image registration, segmentation, and pattern recognition. Over the last decades, most work focus on the removal of additive white Gaussian noise (AWGN) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. Though being effective at Gaussian noise removal, these methods would possibly fail at other types of noise, especially noise in real images, since noise in real images are much more complex than Gaussian [17].

Besides, these methods lack the ability of noise estimation, which makes them impposible for blind image denoising, without the help of other noise estimation algorithm.

Even though they can be adapted for real image denoising with the help of some noise estimation algorithms, these methods do not have the ability to estimate the key parameter of the noise level, which is a key parameter for the performance of image denoising. In order to do real image denoising, these methods require noise estimation as a preprocessing. This make a two-stage framework: a thorough noise estimation method followed by an adapted denoising method. Under this framework, many representative methods achieve good performance in real image denoising [13, 14, 15, 16, 17, 18, 19].

In this perspective, blind image denoising, in which the noise and the level are not known for methods is still a challenge problem. During the image denoising process, these methods need to know the variance of Gaussian noise, which is a key parameter for Gaussian noise removal, in the testing image.

The two limitations indicate that the real image denoising is still a challenging problem.

The noise model in real images are commonly assumed to be Gaussian [13, 14, 15], mixed Gaussian and Laplacian [16], Mixture of Gaussians [17, 18], and tensor Gaussian across the RGB channels [19]. Although Gaussian assumption is commonly made in image denoising, it is inflexible for more complex noise in real images, as suggested in [15, 17]. However, as pointed out in [19], the real camera noise is far from being Gaussian, signal dependent, and usually hard to estimate. Although the Mixture of Gaussian model can be used to approximate any continuous distribution [22], estimating its parameters is time consuming through the commonly employed Expectation-Maximization (EM) algorithm [21] or variational inference technique [22]. On the other hand, the method [16] needs tune specific parameters for different types of noise. This makes the proposed method not a strictly real image denoising. Hence, it is still desirable to design an robust and effective model for real image denoising, which bases on few assumptions and can deal with unknown noise without any parameter tuning procedure. The in-camera imaging pipeline includes image demosaicing, white balance and color space transform, gamut mapping, tone mapping, and JPEG comression. Finally, the major noise in the real im-

age can be categorized into five different types: fixed pattern noise, dark current noise, short noise, amlifier noise, and quantization noise [20].

In this paper, we avoid the challenge problem of noise estimation. We therefore propose a cross domain synthesis solution for real image denoising. In fact, we propose a noval double semi-couplde dictionary learning algorithm for real image denoising problem. In the training stage, given the training patches (niosy ones and clean ones), the dictionaries and coefficients of both the clean and noisy patches.The mapping between the clean and noisy coefficients matrices are also learned in our model. Given a real noisy image, we first extract overlapping patches from it. Then we obtain the clean coefficient from an optimization framework similar to the training stage. The recovered patches are reconstructed by the obtained coefficients on corresponding clean dictionary atoms. We perform comprehensive experiments on real noisy images from multiple different CMOS or CCS sensors. The results demonstrate that our method achieves comparable or even better denoising performance (PSNR, SSIM, and visual quality) on most real noisy images. This reveals that the proposed method has the substantial effect of cross domain image synthesis framework for real image denoising task.

1.1. Our Contributions

To summerize, the contributions of this paper are as follows:

- To the best of our knowledge, we are among the first attempts for real image denoising which regard image denoising as a cross domain transfer problem.

- We also propose a new coupled dictionary learning framework for image restoration problems.

- We demonstrate that our method achieves the state-of-the-art performance on real image denoising problem, both on objective and subjective measurements.

- We construct paired dataset by transforming the unpaired dataset via k-Nearest Neighbor algorithm [?].

- We introduce the Gating Network to speed up the model selection and overall testing speed.

2. Related Work

2.1. Couple dictionary learning

Coupled dictionary learning (CDL) is frequently used in cross-style image synthesis problems such as image super-resolution. CDL assumes that the source and target styles of image have close relationships. CDL aims at learning a pair of dictionaries as well as the relationships between the two cross-domain image styles. Hence, the information

from the source image style can be applied to synthesize the image at the target style. The relationships are often assumed to be identical mapping (coupled) [23], linear mapping (semi-coupled) [24]. Yang et al. [23] assumed that LR image patches have the same sparse representations as their HR versions do, and proposed a joint dictionary learning model for SR using concatenated HR/LR image features. They later imposed relaxed constraints on the observed dictionary/coefficient pairs across image domains for improved performance. Wang et al. [24] further proposed a semi-coupled dictionary learning (SCDL) scheme by advancing a linear mapping for cross-domain image sparse representation. Their method has been successfully applied to applications of image SR and cross-style synthesis.

2.2. Real Image Denoising

To the best of our knowledge, the study of real image denoising can be dated back to the BLS-GSM model [25], in which Portilla et al. proposed to use scale mixture of Gaussian in overcomplete oriented pyramids to estimate the latent clean images. In [13], Portilla proposed to use a correlated Gaussian model for noise estimation of each wavelet subband. Based on the robust statistics theory [?], the work of Rabie [14] modeled the noisy pixels as outliers, which could be removed via Lorentzian robust estimator. In [15], Liu et al. proposed to use 'noise level function' (NLF) to estimate the noise and then use Gaussian conditional random field to obtain the latent clean image. Recently, Gong et al. proposed an optimization based method [16], which models the data fitting term by weighted sum of ℓ_1 and ℓ_2 norms and the regularization term by sparsity prior in the wavelet transform domain. Later, Lebrun el al. proposed a multiscale denoising algorithm called 'Noise Clinic' [17] for real image denoising task. This method generalizes the NL-Bayes [26] to deal with signal, scale, and frequency dependent noise. Recently, Zhu et al. proposed a Bayesian model [18] which approximates the noise via Mixture of Gaussian (MoG) model [22]. The clean image is recovered from the noisy image by the proposed Low Rank MoG filter (LR-MoG). However, noise level estimation is already a challenging problem and denoising methods are quite sensitive to this parameter. Moreover, these methods are based on shrinkage models that are too simple to reflect reality, which results in over-smoothing of important structures such as small-scale text and textures.

3. Double Semi-Couple Dictionary Learning

In this section, we first formulate the real image denoising problem from the perspective of learning based model and then provide the optimization for the problem.

3.1. Problem Formulation

For real image denoising, we first collect clean natural images and real noisy images for training. Assume the $\mathbf{X}$ and $\mathbf{Y}$ are unpaired clean image patches and real noisy patches. Let the $\mathbf{X} = \mathbf{D}_x\mathbf{A}_x$ and $\mathbf{Y} = \mathbf{D}_y\mathbf{A}_y + \mathbf{V}_y$, where $\mathbf{V}_y$ is the real noise of which we don't know the distribution.

$$\min_{\mathbf{D}_x,\mathbf{D}_y,\mathbf{A}_x,\mathbf{A}_y} E_{data}(\mathbf{X},\mathbf{D}_x,\mathbf{A}_x) + E_{data}(\mathbf{Y},\mathbf{D}_y,\mathbf{A}_y,\mathbf{V}_y)$$
$$+ E_{map}(f_1(\mathbf{A}_x),f_2(\mathbf{A}_y)) + E_{reg}(\mathbf{A}_x,\mathbf{A}_y,f_1,f_2,\mathbf{D}_y,\mathbf{V}_y) \quad (1)$$

This framework doesn't need noise modeling and estimation. However, we still model the noise by $\mathbf{V}_y$ for visualization what we have removed during training. The regularization of the noise by $\|\mathbf{V}_y\|_p^p$ can be flexible, that we can penalize it by Frobenius norm, ℓ_1 norm, or any other norms. We employ Frobenius norm here for modeling simplicity. To model the relationship between the representational coefficients, we propose to use two inversible mapping function f_1 and f_2. To measure the error, we employ a penalty function F.

$$\min_{\mathbf{D}_x,\mathbf{D}_y,\mathbf{A}_x,\mathbf{A}_y,\mathbf{U}_x,\mathbf{U}_y,\mathbf{V}_y} \|\mathbf{X} - \mathbf{D}_x\mathbf{A}_x\|_F^2$$
$$+ \|\mathbf{Y} - \mathbf{D}_y\mathbf{A}_y - \mathbf{V}_y\|_F^2 + \alpha F(f_1(\mathbf{A}_x),f_2(\mathbf{A}_y))$$
$$+ \beta_x\|\mathbf{A}_x\|_1 + \beta_y\|\mathbf{A}_y\|_1 + \gamma_y\|\mathbf{V}_y\|_p^p \quad (2)$$
$$\text{s.t.} \quad \|\mathbf{d}_{x,i}\|_2 = 1, \|\mathbf{d}_{y,i}\|_2 = 1, \forall i.$$

Here, we want to discuss more on the mapping functions f_1, f_2 and the measure function F. The mapping function can be linear or nonlinear transformations. The linear function can be defined as a mapping matrix $f_1(\mathbf{A}_x) = \mathbf{U}_x\mathbf{A}_x$ and $f_2(\mathbf{A}_y) = \mathbf{U}_y\mathbf{A}_y$. The corresponding penalty terms on the mapping matrices are $\|\mathbf{U}_x\|_F^2$ and $\|\mathbf{U}_y\|_F^2$. The nonlinear function can be defined as sigmoid function $f_1(\mathbf{A}_x) = 1/(1+\exp\{-\mathbf{A}_y\})$. We can also employ "first-linear-then-nonlinear" or "first-nonlinear-then-linear" strategies. Here, we don't have explict penalty terms for the nonlinear mapping functions. The derivatives of the nonlinear case also need further discussions since it is not easy to obtain closed-form solutions with sigmoid functions. In this paper, we utilize linear transformation matrices as the mapping functions f_1 and f_2. The measure penalty function is simply defined by Frobenius norm. Hence, the term is defined as $\|\mathbf{U}_x\mathbf{A}_x - \mathbf{U}_y\mathbf{A}_y\|_F^2$. However, this would generate a trivial solution of $\mathbf{U}_x = \mathbf{U}_y = \mathbf{0}$. In order to avoid this case, we propose to use the inverse of the mapping matrices, i.e., $\mathbf{U}_x^{-1}$ and $\mathbf{U}_y^{-1}$.

In summary, we propose a Doubly Inversible and Semi-Coupled Dictionary Learing (DISCDL) model to learn the dictionaries and mapping functions between real noisy im-

ages and latent clean natural images.

$$\min_{\mathbf{D}_x,\mathbf{D}_y,\mathbf{A}_x,\mathbf{A}_y,\mathbf{U}_x,\mathbf{U}_y,\mathbf{V}_y} \|\mathbf{X} - \mathbf{D}_x\mathbf{A}_x\|_F^2$$
$$+ \|\mathbf{Y} - \mathbf{D}_y\mathbf{A}_y - \mathbf{V}_y\|_F^2 + \alpha\|\mathbf{U}_x^{-1}\mathbf{A}_x - \mathbf{U}_y^{-1}\mathbf{A}_y\|_F^2$$
$$+ \beta_x\|\mathbf{A}_x\|_1 + \beta_y\|\mathbf{A}_y\|_1 + \gamma_y\|\mathbf{V}_y\|_p^p \quad (3)$$
$$+ \lambda_x\|\mathbf{U}_x^{-1}\|_F^2 + \lambda_y\|\mathbf{U}_y^{-1}\|_F^2$$
$$\text{s.t.} \quad \|\mathbf{d}_{x,i}\|_2 = 1, \|\mathbf{d}_{y,i}\|_2 = 1, \forall i.$$

This model has three major differences when compared with SCDL model.

- We use a matrix $\mathbf{V}_y$ to model the noise, and we don't set any prior distribution on it. This term can help us visualize the noise we learned from the data, i.e., the real noisy images. This make our model fully data-driven. Since our assumption (we have no assumption at all) on noise is more flexible than others', the noise we obtain in our model can be more accurate than other statistical models such as Gaussian or Mixture of Gaussians. Besides, it is time-consuming to fit the noise model from the online data.

- We use two inversible matrices as the mapping transformations between the coefficients of the real noisy patches and the latent clean patches. This makes our model more flexible than SCDL in which the mapping matrix not explictly inversible. Besides, the SCDL can only transform LR images into HG images while our model can transform two different image styles in both direction.

- The constraints on dictionary atoms in our model is strictly $\|\mathbf{d}_{x,i}\|_2 = 1, \|\mathbf{d}_{y,i}\|_2 = 1$ while the CDL model and SCDL model are $\|\mathbf{d}_{x,i}\|_2 \leq 1, \|\mathbf{d}_{y,i}\|_2 \leq 1$. This makes our model more robust on the dictionary learning since both the dictionary atoms and sparse coefficients are interacted with each other. The ≤ 1 constraints would like to make the coefficients larger and dictionary atoms smaller or even vanish. However, in the training stage, we care more about the dictionary atoms and would rather ignore the sparse coefficients.

3.2. Model Optimization

While the objective function in (3) is not convex, it is convex with each varible when other variables are fixed. We employ alternating direction method of multipliers (ADMM) algorithm here. Specifically, we divide the objective function into four sub-problems: 1) updating the sparse coefficients $\mathbf{A}_x, \mathbf{A}_y$; 2) updating the normalized dictionaries $\mathbf{D}_x, \mathbf{D}_y$; 3) updating the noise matrix $\mathbf{V}_y$; 4) updating the mapping matirces $\mathbf{U}_x, \mathbf{U}_y$. We discuss the four steps as follows.

3.2.1 Updating $\mathbf{A}_x$ and $\mathbf{A}_y$

$$\min_{\mathbf{A}_x} \|\mathbf{X} - \mathbf{D}_x\mathbf{A}_x\|_F^2 + \alpha\|\mathbf{U}_x^{-1}\mathbf{A}_x - \mathbf{U}_y^{-1}\mathbf{A}_y\|_F^2 + \beta_x\|\mathbf{A}_x\|_1, \quad (4)$$

$$\min_{\mathbf{A}_y} \|\mathbf{Y} - \mathbf{D}_y\mathbf{A}_y - \mathbf{V}_y\|_F^2$$
$$+\alpha\|\mathbf{U}_x^{-1}\mathbf{A}_x - \mathbf{U}_y^{-1}\mathbf{A}_y\|_F^2 + \beta_y\|\mathbf{A}_y\|_1. \quad (5)$$

Take $\mathbf{A}_x$ as an example, the first and second terms above can be combined to form a new optimization problems as follows:

$$\min_{\mathbf{A}_x} \|\widetilde{\mathbf{X}} - \widetilde{\mathbf{D}}_x\mathbf{A}_x\|_F^2 + \beta_x\|\mathbf{A}_x\|_1, \quad (6)$$

where $\widetilde{\mathbf{X}} = \begin{pmatrix} \mathbf{X} \\ \sqrt{\alpha}\mathbf{U}_y^{-1}\mathbf{A}_y \end{pmatrix}$ and $\widetilde{\mathbf{D}} = \begin{pmatrix} \mathbf{D}_x \\ \sqrt{\alpha}\mathbf{U}_x^{-1} \end{pmatrix}$. For $\mathbf{A}_y$, it is similar with $\mathbf{A}_x$.

$$\min_{\mathbf{A}_y} \|\widetilde{\mathbf{Y}} - \widetilde{\mathbf{D}}_y\mathbf{A}_y\|_F^2 + \beta_y\|\mathbf{A}_y\|_1, \quad (7)$$

where $\widetilde{\mathbf{Y}} = \begin{pmatrix} \mathbf{Y} - \mathbf{V}_y \\ \sqrt{\alpha}\mathbf{U}_x^{-1}\mathbf{A}_x \end{pmatrix}$ and $\widetilde{\mathbf{D}} = \begin{pmatrix} \mathbf{D}_y \\ \sqrt{\alpha}\mathbf{U}_y^{-1} \end{pmatrix}$. These simplified versions have the exactly same formulation as standard sparse coding and can be simply solved by tools such as SPAMS.

The $\mathbf{U}_x^{-1}$ and $\mathbf{U}_y^{-1}$ are invertible. This will be discussed in subsection "Updating U".

3.2.2 Updating $\mathbf{D}_x$ and $\mathbf{D}_y$

$$\min_{\mathbf{D}_x} \|\mathbf{X} - \mathbf{D}_x\mathbf{A}_x\|_F^2 \quad \text{s.t.} \quad \|\mathbf{d}_{x,i}\|_2 = 1, \forall i. \quad (8)$$

$$\min_{\mathbf{D}_y} \|\mathbf{Y} - \mathbf{D}_y\mathbf{A}_y - \mathbf{V}_y\|_F^2 \quad \text{s.t.} \quad \|\mathbf{d}_{y,i}\|_2 = 1, \forall i. \quad (9)$$

These two are quadraically constrained quadratic program (QCQP) problem and can be solved by Lagrange dual techniques.

3.2.3 Updating $\mathbf{V}_y$

The noise matrix is initialized as a zero matirx and updated by solving the following probelm:

$$\min_{\mathbf{V}_y} \|\mathbf{Y} - \mathbf{D}_y\mathbf{A}_y - \mathbf{V}_y\|_F^2 + \gamma_y\|\mathbf{V}_y\|_F^2 \quad (10)$$

This is a ridge regression problem. We can obtain the analytical solution of $\mathbf{V}_y$ by

$$\mathbf{V}_y = (\mathbf{Y} - \mathbf{D}_y\mathbf{A}_y)/(1 + \gamma_y). \quad (11)$$

3.2.4 Updating $\mathbf{U}_x$ and $\mathbf{U}_y$

$$\min_{\mathbf{U}_x^{-1}} \alpha\|\mathbf{U}_y^{-1}\mathbf{A}_y - \mathbf{U}_x^{-1}\mathbf{A}_x\|_F^2 + \lambda_x\|\mathbf{U}_x^{-1}\|_F^2$$
$$\min_{\mathbf{U}_y^{-1}} \alpha\|\mathbf{U}_x^{-1}\mathbf{A}_x - \mathbf{U}_y^{-1}\mathbf{A}_y\|_F^2 + \lambda_y\|\mathbf{U}_y^{-1}\|_F^2 \quad (12)$$

The above problems are also ridge regression problems and have analytical solutions of $\mathbf{U}_x$ and $\mathbf{U}_y$ as follows:

$$\mathbf{U}_x^{-1} = \mathbf{U}_y^{-1}\mathbf{A}_y\mathbf{A}_x^T(\mathbf{A}_x\mathbf{A}_x^T + (\gamma_x/\alpha)\mathbf{I})^{-1}$$
$$\mathbf{U}_y^{-1} = \mathbf{U}_x^{-1}\mathbf{A}_x\mathbf{A}_y^T(\mathbf{A}_y\mathbf{A}_y^T + (\gamma_y/\alpha)\mathbf{I})^{-1} \quad (13)$$

Here, we verify that $\mathbf{U}_x^{-1}$ and $\mathbf{U}_y^{-1}$ are inversible. The $\mathbf{U}_x^{-1}$ and $\mathbf{U}_y^{-1}$ are both initialized as an identity matrix, of suitable dimension, which is inversible. That is, we have $\mathbf{U}_y^{(0)-1} = \mathbf{I}$ when we compute $\mathbf{U}_x^{-1}$. If $\mathbf{A}_y\mathbf{A}_x^T$ is inversible, then $\mathbf{U}_x^{-1}$ is inversible. In fact, we have $\mathbf{A}_y, \mathbf{A}_x \in \mathbb{R}^{d \times N}$. d is the dimension of the sample. For a patch of size 8×8, $d = 64$. The N is the number of samples in the training data. Remember that we have much more samples when compared to the dimension of patches, that is $N \gg d$. It is less likely that $\mathbf{A}_y\mathbf{A}_x^T \in \mathbb{R}^{d \times d}$ has a rank structure lower than d. In other words, $\mathbf{A}_y\mathbf{A}_x^T \in \mathbb{R}^{d \times d}$ is less likely to be singular if we have enough training data. The experiments also confirm our conjecture. Besides, we can also add small disburtation to guarantee that $\mathbf{A}_y\mathbf{A}_x^T \in \mathbb{R}^{d \times d}$ is inversible.

Once $\mathbf{U}_x^{-1}$ is inversible, we can also verify that $\mathbf{U}_y^{-1}$ is inversible in a similar way.

3.3. Real Image Denoising

Two methods:
The first one is that

$$\min_{\mathbf{a}_{x,i},\mathbf{a}_{y,i}} \|\mathbf{x}_i - \mathbf{D}_x\mathbf{a}_{x,i}\|_2^2 + \|\mathbf{y}_i - \mathbf{D}_y\mathbf{a}_{y,i} - \mathbf{v}_{y,i}\|_2^2$$
$$+\alpha\|\mathbf{U}_x^{-1}\mathbf{a}_{x,i} - \mathbf{U}_y^{-1}\mathbf{a}_{y,i}\|_2^2 \quad (14)$$
$$+\beta_x\|\mathbf{a}_{x,i}\|_1 + \beta_y\|\mathbf{a}_{y,i}\|_1 + \gamma_y\|\mathbf{v}_{y,i}\|_1$$

and finally we get $\widehat{\mathbf{x}}_i = \mathbf{D}_x\widehat{\mathbf{a}}_{x,i}$.

The second one is to solve

$$\min_{\mathbf{a}_{y,i},\mathbf{v}_{y,i}} \|\mathbf{y}_i - \mathbf{D}_y\mathbf{a}_{y,i} - \mathbf{v}_{y,i}\|_2^2 + \alpha\|\mathbf{U}_x^{-1}\mathbf{a}_{x,i} - \mathbf{U}_y^{-1}\mathbf{a}_{y,i}\|_2^2$$
$$+\beta_y\|\mathbf{a}_{y,i}\|_1 + \gamma_y\|\mathbf{v}_{y,i}\|_1. \quad (15)$$

Once we get $\widehat{\mathbf{a}}_{y,i}$ from $\mathbf{y}_i$, $\widehat{\mathbf{a}}_{x,i} \approx \mathbf{U}_x\mathbf{U}_y^{-1}\widehat{\mathbf{a}}_{y,i}$ and $\widehat{\mathbf{x}}_i \approx \mathbf{D}_x\widehat{\mathbf{a}}_{x,i}$.

Experiments demonstrate that the first method can get better performance than the second one while the second one can get faster speed than the first one.

We can also initialized the solution from the second one.

4

4. The Overall Algorithm

4.1. Pair Sample Construction from Unpaired Samples

In cross style transfer methods such as CDL and SCDL, the authors assume that the two different styles have paired data, i.e., for each data sample in one style, we can find paired data sample in the other style. However, in real world, the data from two different sources may be unpaired. For example, the real noisy images should not have groundtruth clean images of the same scene. The real low-resolution images should not have corresponding high-resolution images in the real world. The real blurry images should not have corresponding clear and high quality images in real world.

To deal with unpaired data, we could collect real noisy images and clean natural images from two different sources. The real noisy images are from the example images (18 images) of the Neat Image website while the clean natural images are from the training set (200 images) of the Berkeley Segmentation Dataset (BSDS500). To make use of the unpaired data samples, we employ searching strategy to construct the training dataset. That is, for each noisy image patch, we utilize the k-Nearest Neighbor (k-NN) algorithm to find the most similar patch in the clean images as the paired groundtruth patch. The similarity is measured by the Euclidean distance (also called squared error or ℓ_2 norm).

4.2. Structual Clustering and Model Selection

In fact, different image structures should have different influences on dictioanry as well as the mapping function. Patches with flat region should have low rank structure within dictionary elements and identity mapping between noisy and latent clean patches. Patches with complex details should have more comprehensive dictionary elements within dictionary elements and more complex mapping function between noisy and clean patches. A single mapping function cannot deal with all these complex relationships. Hence, a structual clustering procedure is needed for complex solution. In this paper, we propose to employ Gaussian Mixture Model to cluster different image patches into different groups and learn dictionary and mapping function for each group.

4.3. Adaptive Iterations of Different Noise Levels

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.

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 0.0001 in our experiments. In fact, most of our testing images will be denoised well in one iteration.

4.4. Efficient Model Selection by Gating Network

In the Gaussian component selection procedure, if we employ the full posterior estimation, the speed is not fast. Our algorithm can be speeded up by introducing the Gating network model.

5. Experiments

We compare with popular software NeatImage which is one of the best denoising software available. All these methods need noise estimation which is vary hard to perform if there is no uniform regions are available in the testing image. The NeatImage will fail to perform automatical parameters settings if there is no uniform regions.

5.1. Parameters

We don't fine tune the parameters both in the training and testing datasets.

5.2. Real Image Denoising

5.3. Comparison on Dataset with GroundTruth

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] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992. 1

[2] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. *CVPR*, pages 60–65, 2005. 1

[3] S. Roth and M. J. Black. Fields of Experts. *International Journal of Computer Vision*, 82(2):205–229, 2009. 1

[4] 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

[5] 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

[6] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. *ICCV*, pages 2272–2279, 2009. 1

[7] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. *ICCV*, pages 479–486, 2011. 1

[8] 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

[9] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. *CVPR*, pages 2862–2869, 2014. 1

[10] 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

[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

[12] 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

[13] J. Portilla. Full blind denoising through noise covariance estimation using gaussian scale mixtures in the wavelet domain. *Image Processing, 2004. ICIP '04. 2004 International Conference on*, 2:1217–1220, 2004. 1, 2

[14] Tamer Rabie. Robust estimation approach for blind denoising. *Image Processing, IEEE Transactions on*, 14(11):1755–1765, 2005. 1, 2

[15] C. Liu, R. Szeliski, S. Bing Kang, C. L. Zitnick, and W. T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):299–314, 2008. 1, 2

[16] Zheng Gong, Zuowei Shen, and Kim-Chuan Toh. Image restoration with mixed or unknown noises. *Multiscale Modeling & Simulation*, 12(2):458–487, 2014. 1, 2

[17] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image blind denoising. *Image Processing, IEEE Transactions on*, 24(10):3149–3161, 2015. 1, 2

[18] Fengyuan Zhu, Guangyong Chen, and Pheng-Ann Heng. From noise modeling to blind image denoising. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 1, 2

[19] Seonghyeon Nam, Youngbae Hwang, Yasuyuki Matsushita, and Seon Joo Kim. A holistic approach to cross-channel image noise modeling and its application to image denoising. *Proc. Computer Vision and Pattern Recognition (CVPR)*, pages 1683–1691, 2016. 1

[20] Yanghai Tsin, Visvanathan Ramesh, and Takeo Kanade. Statistical calibration of ccd imaging process. *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, 1:480–487, 2001. 2

[21] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)*, pages 1–38, 1977. 1

[22] C. M. Bishop. *Pattern recognition and machine learning*. New York: Springer, 2006. 1, 2

[23] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma. Image super-resolution via sparse representation. *IEEE transactions on image processing*, 19(11):2861–2873, 2010. 2

[24] Shenlong Wang, Lei Zhang, Yan Liang, and Quan Pan. Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis. *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2216–2223, 2012. 2

[25] 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. 2

[26] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3):1665–1688, 2013. 2