ICCV
#572

ICCV
#572

ICCV 2017 Submission #572. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

Multi-channel weighted nuclear norm minimization for color image denoising

Anonymous ICCV submission

Paper ID 572

Abstract

Motivated by the weighted Orthogonal Procrustes Problem, we propose a noval weighted Frobenious norm based weighted sparse coding model for non-Gaussian error modeling. We solve this model in an alternative manner. Updating of each variable has closed-form solutions and the overall model converges to a stationary point. The proposed model is applied in real image denoising problem and extensive experiments demonstrate that the proposed model can much better performance (over 1.0dB improvement on PSNR) than state-of-the-art image denoising methods, including some excellant commercial software. The noval weighted Frobeniius norm can perfectly fit the non-Gaussian property of real noise.

1. Introduction

Image denoising is an important step in enhance the quality of images in computer vision systems. It aims to recover the latent clean image $\mathbf{x}$ from the observed noisy version $\mathbf{y} = \mathbf{x} + \mathbf{n}$, where $\mathbf{n}$ is often assumed to be additive white Gaussian noise. Most denoising methods [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] are designed for grayscale images, and other color image denoising methods [14] treat equally the R, G, B channels in color images. However, in many computer vision tasks, the multiple channels in natural images being processed often exhibit distinct properties, e.g., contain different noise levels. For example, the noise levels among the R, G, B channels are different in real noisy images due to the on board processing in in-camera imaging pipelines [15]. The This is caused by the color demosaicking during the transformation from raw data to RGB images in the standard in-camera imaging pipeline. Usually, the G channel contains the least noise levels among the three channels. Hence, in order to deal with each channel more effectively, different noise levels should be plugged into different channels for color image denoising.

The non-local self similarity (NSS) property of images has been extensively employed in image restoration tasks such as denoising [1, 2, 3, 4, 5, 7, 8]. Among these methods, the weighted nuclear norm minimization (WNNM) model has achieved the state-of-the-art performance on denoising the additive white Gaussian noise (AWGN) in grayscale images. Though among the most effective methods, how to extend the single channel WNNM model to handle multi-channel images such as the real-world color images is still an open problem. Of course the WNNM method can be applied to denoising color images by processing each channel separately, its performance would be largely inferior than jointly processing the RGB channels by concatenating the RGB values into a single vector [14]. Besides, the searching of non-local similar patches would be unstable due to the seperate processing of the RGB images and hence the power of the NSS would be largely reduced. This would also limit the performance of not only WNNM but also other NSS based methods [2, 3, 4, 5, 7]. This fact is also evaluated by our experiments on color image denoising task.

In this paper, we proposed to solve the multi-channel weighted nuclear norm minimization model to perform image denoising on color images. The original WNNM model has closed-form solutions under the weighted nulcear norm proximal operator (WNNP). However, if we add a weighting matrix $\mathbf{W}$ to the left of the data term, the resulting multi-channel WNNM model no longer has the nice property of closed-form solutions. This makes the problem more chanllging. To solve this problem, we formulate the proposed multi-channel WNNM problem into a linearly constrained non-convex program with an augmented variable. It is also not directly solvable due to the non-convexity of the existance of the weighted nuclear norm. Note that the reformulated model contains two variables with linear constraint. This can be solved by employing the alternating direction method of multipliers (ADMM) [16]. For the variable $\mathbf{Z}$, it is the original weighted nuclear norm minimization problem and can be solved with closed-form solution [8, 17]. For the variable $\mathbf{X}$, it is a standard least squares problem and we can also obatin its closed-form solution. The convergency of the proposed method is also given to guarantee a rational termination of the our model.

2. Related Work

2.1. Nuclear Norm Minimization

As the tightest convex surrogate function of the matrix rank minimization [18, 19], the nuclear norm minimization (NNM) problem has been extensive studied in low rank matrix approximation (LRMA) [20, 21, 22, 23]. A standard nuclear norm minimization problem is as follows:

$$\hat{\mathbf{X}} = \arg\min_X \|\mathbf{Y} - \mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_*. \tag{1}$$

This NNM problem has closed-form solution by soft-thresholding the singular values of the matrix $\mathbf{Y}$ as

$$\hat{\mathbf{X}} = \mathbf{U}\mathcal{S}_{\frac{\lambda}{2}}(\mathbf{\Sigma})\mathbf{V}^\top \tag{2}$$

where $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ is the singular value decomposition (a.k.a. Eckart-Young Decomposition [24]) of $\mathbf{Y}$ and $\mathcal{S}_\tau(\bullet)$ is the soft-thresholding function with parameter $\tau > 0$:

$$\mathcal{S}_\tau(\mathbf{\Sigma}_{ii}) = \max(\mathbf{\Sigma}_{ii} - \tau, 0) \tag{3}$$

One limitation of the original NNM model is that it treats all the singular values equally but ignore the different importance of them. To make the NNM mmodel more flexible at processing sigular values, it has been extended to the truncated nuclear norm minimization model [25], the partial sum minimization of singular values [26], and the weighted nuclear norm minimization (WNNM) model [8], etc. Among these models, the WNNM model has been applied on grayscale image denoising problem with highly effective performance. This model adds weights to each singular values and the problem is:

$$\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \|\mathbf{X}\|_{\boldsymbol{w},*} \tag{4}$$

is firstly proposed for grayscale image denoising problem, where $\|\mathbf{X}\|_{\boldsymbol{w},*} = \sum_i w_i \sigma_i(\mathbf{X})$ is the weighted nuclear norm of matrix $\mathbf{X}$ and $\boldsymbol{w} = [w_1, ..., w_n]^\top, w_i \geq 0$ is the weight vector. According to the Remark 1 of [27], the problem (4) has closed-form solution if the weights are in a non-decreasing order

Though having achieved excellent performance on grayscale image denoising task, the WNNM method could not be applied on color image denoising in a direct manner. Of course we can apply the WNNM on each channel seperately, but it has been studied that this manner would get inferior performance when compared to the power of this model on grayscale images. In this paper, we would add a weighting matrix to the WNNM model and naturally extend it to deal with color images and maintain its powerful ability on exploring the non-local self similarity property of the natural images. The proposed multi-channel WNNM model can be solved via the famous ADMM [16] algorithm. For each variable, we can derive a closed-form solution and hence the overall model can be solved in an efficient way. Besides, we analysis the convergence results of the proposed model.

2.2. Color Image Denoising

Numerous methods [1, 2, 4, 5, 7, 8, 9, 10, 11, 13] has been proposed for grayscale image denoising task. These methods could be trivially employed to denoise each channel in color images separately. For example, the CBM3D [6] first transform the RGB image into luminance-chrominance color space (e.g., YCbCr) and perform BM3D [5] for each channels separately with the patches only being grouped in luminance channel. However, as pointed out in [14], the performance would be largely reduced when compared to the performance of these methods on grayscale images. This method [14] performs color image denoising by concatenating the patches in R, G, B channels into a single vector. However, the concatenation would generate false colors and artifacts [14]. The major reason is that the concatenation methods treat multiple channels equally and ignore the different properties among these channels. To better explore the difference and correlation among the three channels in color images, several methods [28, 29, 30, 31, 32, 33, 34] have been proposed to deal with real noisy image denoising task. These methods often firstly estimate the parameters of the assumed noise model (usually Gaussian), and then perform denoising with the estimated noise model. In this paper, we propose a weighting matrix (diagonal) which add differents noise levels on different channels for color image denoising. The proposed multi-channel method are able to deal with the distinct noise property among these channels. Experiments will demonstrate that the proposed model achieves better results than other competing methods for color image denoising.

3. Multi-channel Weighted Nuclear Norm Minimization

3.1. The Problem

Given In order to treat differently each channel in color images, we propose to add a weighting matrix to the original WNNM model and the resulting problem becomes

$$\min_{\mathbf{X}} \|\mathbf{W}(\mathbf{Y} - \mathbf{X})\|_F^2 + \|\mathbf{X}\|_{\boldsymbol{w},*}. \tag{5}$$

where $\mathbf{W}$ is the weighting matrix. For simplisity, we assume $\mathbf{W}$ to be a diagonal matrix.

Unfortunately, the proposed multi-channel WNNM problem cannot be solved in an analytical form. In [27], when the weights on singular values are non-descending, the weighted nuclear norm proximal operator can have global optimum with closed-form solution. However, such property is not valid for the multi-channel WNNM model. The reason is that the weighting matrix $\mathbf{W}$ is added to the matrix $\mathbf{X}$ instead of its singular values. Besides, the elements in $\mathbf{W}$ is not in a non-descending order with respect

to the singular value of $\mathbf{X}$. This makes the proposed model more difficult to optimize than the original WNNM model.

This can be solved by introducing an augmented variable $\mathbf{Z}$, and the above multi-channel WNNM problem is equivalent to a linearly constrained non-convex problem with two variables.

$$\min_{\mathbf{X},\mathbf{Z}} \|\mathbf{W}(\mathbf{Y} - \mathbf{X})\|_F^2 + \|\mathbf{Z}\|_{\boldsymbol{w},*} \quad \text{s.t.} \quad \mathbf{X} = \mathbf{Z}. \quad (6)$$

This is an optimization problem with functions of two variables $\mathbf{X}$ and $\mathbf{Z}$ with linearly constrained condition of $\mathbf{X} = \mathbf{Z}$. In fact, this proplem can be solved by the alternating direction method of multipliers (ADMM) algorithm, which will be introduced in the next subsection.

3.2. Optimization

To solve the above optimization problem, we first derive its augmented Lagrangian function as

$$\mathcal{L}(\mathbf{X},\mathbf{Z},\mathbf{A},\rho) = \|\mathbf{W}(\mathbf{Y} - \mathbf{X})\|_F^2 + \|\mathbf{Z}\|_{\boldsymbol{w},*}$$
$$+ \langle \mathbf{A}, \mathbf{X} - \mathbf{Z} \rangle + \frac{\rho}{2}\|\mathbf{X} - \mathbf{Z}\|_F^2 \quad (7)$$

where $\mathbf{A}$ is the augmented Lagrangian multiplier and $\rho > 0$ is the penalty parameter. After some simple calculations, we can obtain the following equivalent form of the Lagrangian function

$$\mathcal{L}(\mathbf{X},\mathbf{Z},\mathbf{A},\rho) = \|\mathbf{W}(\mathbf{Y} - \mathbf{X})\|_F^2 + \|\mathbf{Z}\|_{\boldsymbol{w},*}$$
$$+ \frac{\rho}{2}\|\mathbf{X} - \mathbf{Z} + \rho^{-1}\mathbf{A}\|_F^2 \quad (8)$$

We initialize the matrix variables $\mathbf{X}_0$, $\mathbf{Z}_0$, and $\mathbf{A}_0$ to be zero matrix of suitable size. Taking derivative of the Lagrangian function $\mathcal{L}$ with respect to the variables $\mathbf{X}$ and $\mathbf{Z}$ and setting the derivative function to be zero, we can alternatively update the iterations of the ADMM algorithm as follows:

(1) Update $\mathbf{X}$ while fixing $\mathbf{Z}$ and $\mathbf{A}$:

$$\mathbf{X}_{k+1} = \arg\min_{\mathbf{X}} \|\mathbf{W}\mathbf{Y} - \mathbf{W}\mathbf{X}\|_F^2 + \frac{\rho_k}{2}\|\mathbf{X} - \mathbf{Z}_k + \rho_k^{-1}\mathbf{A}_k\|_F^2 \quad (9)$$

This is a mixed weighted least square and standard least square problem and we could derive its closed-form solution:

$$\mathbf{X}_{k+1} = (\mathbf{W}^\top\mathbf{W} + \frac{\rho_k}{2}\mathbf{I})^{-1}(\mathbf{W}^\top\mathbf{W}\mathbf{Y} + \frac{\rho_k}{2}\mathbf{Z}_k - \frac{1}{2}\mathbf{A}_k) \quad (10)$$

(2) Update $\mathbf{Z}$ while fixing $\mathbf{X}$ and $\mathbf{A}$:

$$\mathbf{Z}_{k+1} = \arg\min_{\mathbf{Z}} \frac{\rho_k}{2}\|\mathbf{Z} - (\mathbf{X}_{k+1} + \rho_k^{-1}\mathbf{A}_k)\|_F^2 + \|\mathbf{Z}\|_{\boldsymbol{w},*} \quad (11)$$

According to the Theorem 1 in [27], given the $\mathbf{X}_{k+1} + \rho_k^{-1}\mathbf{A}_k = \mathbf{U}_k\boldsymbol{\Sigma}_k\mathbf{V}_k^\top$ be the SVD of $\mathbf{X}_{k+1} + \rho_k^{-1}\mathbf{A}_k$,

where $\boldsymbol{\Sigma}_k = \begin{pmatrix} \text{diag}(\sigma_1, \sigma_2, ..., \sigma_n) \\ \mathbf{0} \end{pmatrix} \in \mathbb{R}^{m \times n}$, then the global optimum of the above problem is $\hat{\mathbf{Z}} = \mathbf{U}_k\hat{\boldsymbol{\Sigma}}_k\mathbf{V}_k^\top$, where $\hat{\boldsymbol{\Sigma}}_k = \begin{pmatrix} \text{diag}(\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_n) \\ \mathbf{0} \end{pmatrix} \in \mathbb{R}^{m \times n}$ and $(\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_n)$ is the solution to the following convex optimization problem:

$$\min_{\hat{\sigma}_1, \hat{\sigma}_2, ..., \hat{\sigma}_n} \sum_{i=1}^n (\sigma_i - \hat{\sigma}_i)^2 + \frac{2w_i}{\rho_k}\hat{\sigma}_i \quad (12)$$
$$\text{s.t.} \quad \hat{\sigma}_1 \geq \hat{\sigma}_2 \geq ... \geq \hat{\sigma}_n \geq 0.$$

According to the Remark 1 in [27], the problem above has closed-form solution

$$\hat{\sigma}_i = \begin{cases} 0 & \text{if } c_2 < 0 \\ \frac{c_1 + \sqrt{c_2}}{2} & \text{if } c_2 \geq 0 \end{cases} \quad (13)$$

where $c_1 = \sigma_i - \epsilon$, $c_2 = (\sigma_i - \epsilon)^2 - \frac{8C}{\rho_k}$ and C is set as $\sqrt{2n}$ by experience in image denoising.

(3) Update $\mathbf{A}$ while fixing $\mathbf{X}$ and $\mathbf{Z}$:

$$\mathbf{A}_{k+1} = \mathbf{A}_k + \rho_k(\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}) \quad (14)$$

(4) Update ρ_k as $\rho_{k+1} = \mu * \rho_k$.

The above 4 alternative updating steps are repeated until the convergence conditions are satisfied or the number of iterations exceeds a preset maximum number, e.g., K_1. The overall algorithm will achieve its convergence conditions when $\|\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F \leq \text{Tol}$, $\|\mathbf{X}_{k+1} - \mathbf{X}_k\|_F \leq \text{Tol}$, and $\|\mathbf{Z}_{k+1} - \mathbf{Z}_k\|_F \leq \text{Tol}$ are simutaniously satisfied, where $\text{Tol} > 0$ is a small tolerance. We summerize the optimization steps in Algorithm 1 (A1). We give a theorem, i.e., Theorem 1, to guarantee the convergence of the proposed Algorithm 1. Note that since the weighted nuclear norm is non-convex in general, we employ an unbounded sequence of $\{\rho_k\}$ here to make sure that the Algorithm 1 is convergent.

Theorem 1. *Assume the weights in $\boldsymbol{w}$ are in a non-descending order, the sequence $\{\mathbf{X}_k\}$, $\{\mathbf{Z}_k\}$, and $\{\mathbf{A}_k\}$ generated in Algorithm 1 satisfy:*

$$(1) \lim_{k \to \infty} \|\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F = 0; \quad (15)$$

$$(2) \lim_{k \to \infty} \|\mathbf{X}_{k+1} - \mathbf{X}_k\|_F = 0; \quad (16)$$

$$(3) \lim_{k \to \infty} \|\mathbf{Z}_{k+1} - \mathbf{Z}_k\|_F = 0. \quad (17)$$

Proof. We give proof sketch here and detailed proof of this theorem can be found in Appendix. We can first proof that the sequence $\{\mathbf{A}_k\}$ generated by Algorithm 1 is upper bounded. Since $\{\rho_k\}$ is unbounded, that is $\lim_{k \to \infty} \rho_k = +\infty$, we can proof that the sequence of

ICCV
#572

ICCV
#572

ICCV 2017 Submission #572. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

A1: Solve Multi-channel WNNM via ADMM

Input: Matrices $\mathbf{Y}$ and $\mathbf{W}$, $\mu > 1$, Tol > 0, $K_1 > 0$;

Initialization: $\mathbf{X}_0 = \mathbf{Z}_0 = \mathbf{A}_0 = \mathbf{0}$, $\rho_0 > 0$, T = False, $k = 0$;

While (T == false) **do**

1. Update $\mathbf{X}_{k+1}$ as
$\mathbf{X}_{k+1} = (\mathbf{W}^\top\mathbf{W} + \frac{\rho_k}{2}\mathbf{I})^{-1}(\mathbf{W}^\top\mathbf{W}\mathbf{Y} + \frac{\rho_k}{2}\mathbf{Z}_k - \frac{1}{2}\mathbf{A}_k)$

2. Update $\mathbf{Z}_{k+1}$ by solving the WNNM problem
$\min_{\mathbf{Z}} \frac{\rho_k}{2}\|\mathbf{Z} - (\mathbf{X}_{k+1} + \rho_k^{-1}\mathbf{A}_k)\|_F^2 + \|\mathbf{Z}\|_{\boldsymbol{w},*}$

3. Update $\mathbf{A}_{k+1}$ as $\mathbf{A}_{k+1} = \mathbf{A}_k + \rho_k(\mathbf{X}_{k+1} - \mathbf{Z}_{k+1})$

4. Update $\rho_{k+1} = \mu * \rho_k$;

5. $k \leftarrow k + 1$;
  if ($\|\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F/\|\mathbf{Z}_{k+1}\|_F <$ Tol) or ($k \geq K_1$)

5. T $\leftarrow$ True
  end if

end while

Output: Matrices $\mathbf{X}$ and $\mathbf{Z}$.

Lagrangian function $\{\mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_k, \rho_k)\}$ is also upper bounded. Hence, both $\{\mathbf{WY} - \mathbf{WX}_k\}$ and $\{\mathbf{Z}_k\}$ are upper bounded. According to Eq. (14), we can proof that $\lim_{k\to\infty}\|\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F = \lim_{k\to\infty}\rho_k^{-1}\|\mathbf{A}_{k+1} - \mathbf{A}_k\|_F = 0$, and (1) is proofed. Then we can proof that $\lim_{k\to\infty}\|\mathbf{X}_{k+1} - \mathbf{X}_k\|_F \leq \lim_{k\to\infty}\|(\mathbf{W}^\top\mathbf{W} + \frac{\rho_k}{2}\mathbf{I})^{-1}(\mathbf{W}^\top\mathbf{W}\mathbf{Y} - \mathbf{W}^\top\mathbf{W}\mathbf{Z}_k - \frac{1}{2}\mathbf{A}_k)\|_F + \rho_k^{-1}\|\mathbf{A}_k - \mathbf{A}_{k-1}\|_F = 0$ and hence (2) is proofed. Then (3) can be proofed by checking that $\lim_{k\to\infty}\|\mathbf{Z}_{k+1} - \mathbf{Z}_k\| \leq \lim_{k\to\infty}\|\boldsymbol{\Sigma}_{k-1} - \mathcal{S}_{\boldsymbol{w}/\rho_{k-1}}(\boldsymbol{\Sigma}_{k-1})\|_F + \|\mathbf{X}_{k+1} - \mathbf{X}_k\|_F + \rho_k^{-1}\|\mathbf{A}_{k-1} + \mathbf{A}_{k+1} - \mathbf{A}_k\|_F = 0$, where $\mathbf{U}_{k-1}\boldsymbol{\Sigma}_{k-1}\mathbf{V}_{k-1}^\top$ is the SVD of the matrix $\mathbf{X}_k + \rho_{k-1}\mathbf{A}_{k-1}$. The proof sketch of Theorem 1 is end. □

4. Multi-channel WNNM For Color Image Denoising

In this section, we apply the proposed multi-channel WNNM model on color image denoising problem. The multi-channel WNNM model can make use of the non-local self similarity property of natural images while treating each channel adaptively. In real-world noisy images, the noise are first emerged in the RAW data, i.e., color filter array (CFA). The major noise generated in real noisy images are due to the discrete nature of light and thermal agitation [], which can be modeled as Poisson and Gaussian distribution, respectively. Since the Poisson distribution can be approximately modeled by Gaussian distribution, the overall noise model in each channel of the color image could be Gaussian. Some methods would first transform the color images into other color spaces such as luminance/chrominance space such as YCbCr or Lab, but this would change the noise distribution in each channel. Hence, in this work we still choose to deal with the RGB channels in color images. Besides, even though if the demosaicing of RAW image generate similar distribution in noise in different channels, the channel-wise scaling in white balance would definitely change the noise distribution in each channel. Thus, the noise in R, G, B channels are definitely different which can be described by different noise levels and structures. According to above analysis, color image denoising is to recover the latent clean image $\mathbf{x}$ from the observed noisy version $\mathbf{y}_c = \mathbf{x}_c + \mathbf{n}_c$, where $c \in \{R, G, B\}$ represent the R, G, B channels in color images and $\mathbf{n}_c$ is the noise in the c channel (assumed to be additive white Gaussian noise).

The patches in color image $\mathbf{y}$ are of size $p \times p \times 3$. For each patch $\mathbf{y}_j$, we search its non-local similar patches in a large area and stack the similar patches column by column. The resulting matrix $\mathbf{Y}_j \in \mathbb{R}^{3p^2 \times n}$, where n is the number of similar patches. The corresponding matrices containing the clean patches and the channel-wise noise are defined as $\mathbf{X}_j$ and $\mathbf{N}_j$, respectively. Since $\mathbf{X}_j$ is made of similar patches, it should be a low rank matrix. And hence the multi-channel WNNM model proposed in this paper can be used here. Compared to the original single channel WNNM model [8] proposed for grayscale image denoising

$$\min_{\mathbf{X}_j} \|\mathbf{W}(\mathbf{Y}_j - \mathbf{X}_j)\|_F^2 + \|\mathbf{X}_j\|_{w,*}. \qquad (18)$$

When the weighting matrix $\mathbf{W} = \frac{1}{\sigma_n^2}\mathbf{I}$, where $\mathbf{I} \in \mathbb{R}^{3p^2 \times 3p^2}$ is the identity matrix, the multi-channel WNNM model will reduce to the WNNM model as a special case. The design of WNNM model also motivate us to consider a similar design of the weighting matrix. In order to deal with color image denoising task, the weighting matrix $\mathbf{W}$ should be modified to be suitable for multi-channel cases. In fact, for the color image denoising task, a holistic try of the weighting matrix $\mathbf{W}$ could be

$$\mathbf{W} = \begin{pmatrix} \sigma_r^{-2}\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_g^{-2}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_b^{-2}\mathbf{I} \end{pmatrix}. \qquad (19)$$

Here, for simplicity, we assume that the noise in different channels are independent to each other. The experimental results have already demonstrated that this simple assumption could already generate the best denoising performance on benchmark real noisy image dataset. In this paper, we did not consider the correlations of noise among different channels, which is the future work of our research line. The determination of the weight vector in weighted nuclear norm is the same as in the WNNM model [27]. We set the weight vector as $w_i^{k+1} = \frac{C}{|\sigma_i(\mathbf{X}_k)| + \epsilon}$.

The multi-channel WNNM is applied to the non-local similar patches of each local patch in the noisy image $\mathbf{y}$. And then all the patches are aggregated together to form the final recovered image $\hat{\mathbf{y}}$. We also perform the denoising

ICCV
#572

ICCV
#572

ICCV 2017 Submission #572. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

A2: Color Image Denoising by Multi-channel WNNM

Input: Noisy image $\mathbf{y}$, noise levels $\{\sigma_r, \sigma_g, \sigma_b\}$;

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

for $k = 1 : K_2$ **do**

1. Set $\mathbf{y}^{(k)} = \hat{\mathbf{x}}^{(k-1)}$;

2. Extracte local patches $\{\mathbf{y}_j\}_{j=1}^N$ from $\mathbf{y}^{(k)}$;

   for each patch $\mathbf{y}_j$ **do**

3.   Search non-local similar patches $\mathbf{Y}_j$;

4.   Estimate $\mathbf{X}_j$ by applying the Algorithm 1 to $\mathbf{Y}_j$;

   end for

5. Aggregate $\{\mathbf{X}_j\}_{j=1}^N$ to form the image $\hat{\mathbf{x}}^{(k)}$;

end for

Output: Denoised image $\hat{\mathbf{x}}^{K_2}$.

---

procedure for several (K_2) iterations to obtain better denoising results. In Algorithm 2 (A2), we summarizes the denoising steps of multi-channel WNNM model on color image denoising.

5. Experiments

We evaluate the proposed method on synthetic noisy images as well as real noisy images. We compare the proposed method with other state-of-the-art denoising algorithms including CBM3D [5, 6], MLP [9], CSF [10], WNNM [8], TNRD [11], "Noise Clinic" [32, 33], and the commercial software Neat Image [34].

In order to compare with the original WNNM method, we not only apply the WNNM method on each channel separately, but also concatenate the RGB values into a single vector to perform denoising to boost the original WNNM method with channel correlations. To demonstrate the effectiveness of the proposed weighting matrix $\mathbf{W}$, we set $\mathbf{W} = \sigma_n \mathbf{I}$, which is reduced to the standard WNNM method. These three types of WNNM are named as WNNM, WNNMc, and WNNMadmm, respectively.

5.1. Experiments on Synthetic Noisy Images

In this section, we compare the proposed method with other competing method in 24 high quality color images from the Kodak PhotoCD Dataset (http://r0k.us/graphics/kodak/), which are shown in Fig. 1. Then we add additive white Gaussian noise with different standard deviations to different channels of the color images. The standard deviations of noise we add to the R, G, B channels of the 24 clean images are 40, 20, 30, respectively. We set the patch size as $p = 6$, the number of non-local similar patches as $n = 70$, the window size for searching similar patches as $W = 20$. For the proposed multi-channel WNNM model, we set the regularization parameter as $\lambda = 4$, the penalty parameter as $\rho = 6$, the $\mu = 1.1$, the number of iterations in Algorithm 1 as $K_1 = 10$, the number of iterations in Algorithm 2 as $K_2 = 8$. For fair
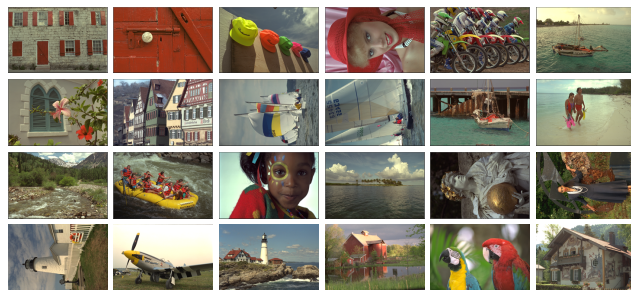

Figure 1. The 24 high quality color images from the Kodak PhotoCD Dataset.

comparison, we employ the same parameter settings for the WNNM, WNNMc, WNNMadmm methods.

5.2. Experiments on Real Noisy Images

In the second part, we compare the proposed method with other competing methods on the 15 real noisy images, , which are shown in Fig. 2, with "ground truth" clean images [28]. The noisy images were collected under controlled indoor environment. Each scene was shot 500 times under the same camera and camera setting. The mean image of the 500 shots is roughly taken as the "ground truth", with which the PSNR can be computed. Since the image size is very large (about 7000×5000) and the scenes of this dataset share repetitive contents, the authors of [28] cropped 15 smaller images (of size 512×512) to perform experiments.

We firstly perform quantitative comparison on the 15 cropped images used in [28]. The PSNR results of CBM3D [5], WNNM [8], MLP [9], CSF [10], TNRD [11], NC [32, 33], NI [34] and CC [28] are listed in Table 2 (The results of CC are copied from the original paper [28]). The best and second best PSNR results of each image are highlighted in red and blue, respectively. One can see that on 9 out of the 15 images, our method achieves the best PSNR values. CC achieves the best PSNR on 3 of the 15 images. It should be noted that in the CC method, a specific model is trained for each camera and camera setting, while our method uses the same model for all images. On average, our proposed method has 0.28dB PSNR improvements over [28] and much higher PSNR gains over other competing methods. Fig. **??** shows the denoised images of a scene captured by Canon 5D Mark 3 at ISO = 3200. We can see that CBM3D, WNNM, NC, NI and CC would either remain noise or generate artifacts, while MLP, TNRD over-smooth much the image. By using the external prior guided internal priors, our proposed method preserves edges and textures better than other methods, leading to visually pleasant outputs. More visual comparisons can be found in the supplementary file.

ICCV
#572

ICCV
#572

ICCV 2017 Submission #572. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

Table 1. PSNR(dB) results of different denoising algorithms on 20 natural images.

Image#		CBM3D	MLP	CSF	TNRD	Noise Clinic	Neat Image	WNNM	WNNMc	WNNMadmm	Multi-channel WNNM
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
Average											

Header row spans: $\sigma_r = 40, \sigma_g = 20, \sigma_b = 30$

Table 2. PSNR(dB) results of different methods on 15 cropped real noisy images used in [28].

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

6. Conclusion

The noise in real-world noisy images is much more complex than additive white Gaussian noise and hard to be modeled by simple analytical distributions such as Gaussian or mixture of Gaussians, making the real noisy image denoising problem very challenging. In this paper, we proposed a simple yet effective solution for real noisy image denoising without explicitly assuming certain noise models. Specifically, we firstly learned image priors from external clean images, and then employed the learned external priors to guide the learning of internal priors from the given noisy image. The learned hybrid priors are utilized for real noisy image denoising. Experiments on two real noisy image datasets, where the images were captured under indoor or outdoor lighting conditions by different types of cameras and camera settings, demonstrated that our proposed method achieved much better performance than state-of-the-art image denoising methods.
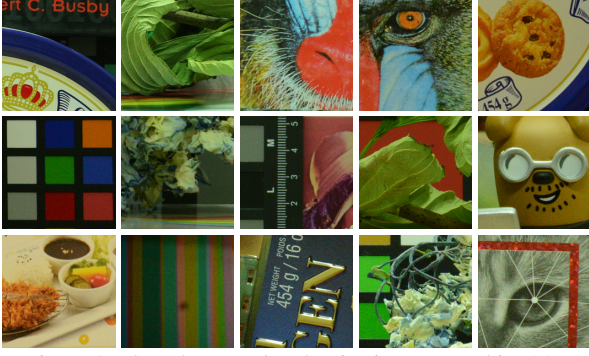
ICCV
#572

ICCV 2017 Submission #572. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

ICCV
#572

Figure 2. The 15 cropped real noisy images used in [28].

7. A. Proof of Theorem 1.

Proof. 1. Firstly, we proof that the sequence $\{\mathbf{A}_k\}$ generated by Algorithm 1 is upper bounded. Let $\mathbf{X}_{k+1} + \rho_k^{-1}\mathbf{A}_k = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^\top$ be its SVD in the $(k+1)$-th iteration. According to Corollary 1 of [27], we can have the SVD of $\mathbf{Z}_{k+1}$ as $\mathbf{Z}_{k+1} = \mathbf{U}_k\hat{\mathbf{\Sigma}}_k\mathbf{V}_k^\top = \mathbf{U}_k\mathcal{S}_{\frac{w}{\rho_k}}(\mathbf{\Sigma}_k)\mathbf{V}_k^\top$. Then we have

$$\|\mathbf{A}_{k+1}\|_F = \|\mathbf{A}_k + \rho_k(\mathbf{X}_{k+1} - \mathbf{Z}_{k+1})\|_F \tag{20}$$

$$= \rho_k\|\rho_k^{-1}\mathbf{A}_k + \mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F \tag{21}$$

$$= \rho_k\|\mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^\top - \mathbf{U}_k\mathcal{S}_{\frac{w}{\rho_k}}(\mathbf{\Sigma}_k)\mathbf{V}_k^\top\|_F \tag{22}$$

$$= \rho_k\|\mathbf{\Sigma}_k - \mathcal{S}_{\frac{w}{\rho_k}}(\mathbf{\Sigma}_k)\|_F \tag{23}$$

$$= \rho_k\sqrt{\sum_i(\mathbf{\Sigma}_k^{ii} - \mathcal{S}_{\frac{w_i}{\rho_k}}(\mathbf{\Sigma}_k^{ii}))^2} \tag{24}$$

$$\leq \rho_k\sqrt{\sum_i(\frac{w_i}{\rho_k})^2} = \sqrt{\sum_i w_i^2}. \tag{25}$$

The inequality above can be proofed as follows: given the diagonal matrix $\mathbf{\Sigma}_k$, we define $\mathbf{\Sigma}_k^{ii}$ as the i-th element of $\mathbf{\Sigma}_k^{ii}$. If $\mathbf{\Sigma}_k^{ii} \geq \frac{w_i}{\rho_k}$, we have $\mathcal{S}_{\frac{w_i}{\rho_k}}(\mathbf{\Sigma}_k^{ii}) = \mathbf{\Sigma}_k^{ii} - \frac{w_i}{\rho_k} > 0$. If $\mathbf{\Sigma}_k^{ii} < \frac{w_i}{\rho_k}$, we have $\mathcal{S}_{\frac{w_i}{\rho_k}}(\mathbf{\Sigma}_k^{ii}) = 0 < \mathbf{\Sigma}_k^{ii} + \frac{w_i}{\rho_k}$. After all, we have $|\mathbf{\Sigma}_k^{ii} - \mathcal{S}_{\frac{w_i}{\rho_k}}(\mathbf{\Sigma}_k^{ii})| \leq \frac{w_i}{\rho_k}$ and hence the inequality holds true. Hence, the sequence $\{\mathbf{A}_k\}$ is upper bounded.

2. Secondly, we proof that the sequence of Lagrangian function $\{\mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_k, \rho_k)\}$ is also upper bounded. Since the global optimal solution of $\mathbf{X}$ and $\mathbf{Z}$ in corresponding subproblems, we always have $\mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_k, \rho_k) \leq \mathcal{L}(\mathbf{X}_k, \mathbf{Z}_k, \mathbf{A}_k, \rho_k)$. Based on the updating rule that $\mathbf{A}_{k+1} = \mathbf{A}_k + \rho_k(\mathbf{X}_{k+1} - \mathbf{Z}_{k+1})$, we have $\mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_{k+1}, \rho_{k+1}) = \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_k, \rho_k) + \langle\mathbf{A}_{k+1} - \mathbf{A}_k, \mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\rangle + \frac{\rho_{k+1} - \rho_k}{2}\|\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F^2 = \mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_k, \rho_k) + \frac{\rho_{k+1} + \rho_k}{2\rho_k^2}\|\mathbf{A}_{k+1} - \mathbf{A}_k\|_F^2$. Since the sequence $\{\|\mathbf{A}_k\|\}$ is upper bounded, the sequence $\{\|\mathbf{A}_{k+1} - \mathbf{A}_k\|_F\}$ is also upper bounded. Denote by a the upper bound of $\{\|\mathbf{A}_{k+1} - \mathbf{A}_k\|_F\}$, we have

$\mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_{k+1}, \rho_{k+1}) \leq \mathcal{L}(\mathbf{X}_1, \mathbf{Z}_1, \mathbf{A}_0, \rho_0) + a\sum_{k=0}^{\infty}\frac{\rho_{k+1} + \rho_k}{2\rho_k^2} = \mathcal{L}(\mathbf{X}_1, \mathbf{Z}_1, \mathbf{A}_0, \rho_0) + a\sum_{k=0}^{\infty}\frac{\mu+1}{2\mu^k\rho_0} \leq \mathcal{L}(\mathbf{X}_1, \mathbf{Z}_1, \mathbf{A}_0, \rho_0) + \frac{a}{\rho_0}\sum_{k=0}^{\infty}\frac{1}{\mu^{k-1}}$. The last inequality holds since $\mu + 1 < 2\mu$. Since $\sum_{k=0}^{\infty}\frac{1}{\mu^{k-1}} < \infty$, the sequence of Lagrangian function $\mathcal{L}(\mathbf{X}_{k+1}, \mathbf{Z}_{k+1}, \mathbf{A}_{k+1}, \rho_{k+1})$ is upper bound.

3. Thirdly, we proof that the sequences of $\{\mathbf{X}_k\}$ and $\{\mathbf{Z}_k\}$ are upper bounded. Since $\|\mathbf{W}(\mathbf{Y} - \mathbf{X})\|_F^2 + \|\mathbf{Z}\|_{\boldsymbol{w},*} = \mathcal{L}(\mathbf{X}_k, \mathbf{Z}_k, \mathbf{A}_{k-1}, \rho_{k-1}) - \langle\mathbf{A}_k, \mathbf{X}_k - \mathbf{Z}_k\rangle - \frac{\rho_k}{2}\|\mathbf{X}_k - \mathbf{Z}_k\|_F^2 = \mathcal{L}(\mathbf{X}_k, \mathbf{Z}_k, \mathbf{A}_{k-1}, \rho_{k-1}) + \frac{1}{2\rho_k}(\|\mathbf{A}_{k-1}\|_F^2 - \|\mathbf{A}_k\|_F^2)$. Thus $\{\mathbf{W}(\mathbf{Y} - \mathbf{X}_k)\}$ and $\{\mathbf{Z}_k\}$ are upper bounded, and hence the sequence $\{\mathbf{X}_k\}$ is bounded by Cauchy-Schwarz inequality and triangle inequality. We can obtain that $\lim_{k\to\infty}\|\mathbf{X}_{k+1} - \mathbf{Z}_{k+1}\|_F = \lim_{k\to\infty}\rho_k^{-1}\|\mathbf{A}_{k+1} - \mathbf{A}_k\|_F = 0$ and the equation (1) is proofed.

4. Then we can proof that $\lim_{k\to\infty}\|\mathbf{X}_{k+1} - \mathbf{X}_k\|_F = \lim_{k\to\infty}\|(\mathbf{W}^\top\mathbf{W} + \frac{\rho_k}{2}\mathbf{I})^{-1}(\mathbf{W}^\top\mathbf{W}\mathbf{Y} - \mathbf{W}^\top\mathbf{W}\mathbf{Z}_k - \frac{1}{2}\mathbf{A}_k) - \rho_k^{-1}(\mathbf{A}_k - \mathbf{A}_{k-1})\|_F \leq \lim_{k\to\infty}\|(\mathbf{W}^\top\mathbf{W} + \frac{\rho_k}{2}\mathbf{I})^{-1}(\mathbf{W}^\top\mathbf{W}\mathbf{Y} - \mathbf{W}^\top\mathbf{W}\mathbf{Z}_k - \frac{1}{2}\mathbf{A}_k)\|_F + \rho_k^{-1}\|\mathbf{A}_k - \mathbf{A}_{k-1}\|_F = 0$ and hence (2) is proofed.

5. Then (3) can be proofed by checking that $\lim_{k\to\infty}\|\mathbf{Z}_{k+1} - \mathbf{Z}_k\|_F = \lim_{k\to\infty}\|\mathbf{X}_k + \rho_k^{-1}\mathbf{A}_{k-1} - \mathbf{Z}_k + \mathbf{X}_{k+1} - \mathbf{X}_k + \rho_k^{-1}\mathbf{A}_{k-1} + \rho_k^{-1}\mathbf{A}_k - \rho_k^{-1}\mathbf{A}_{k+1}\|_F \leq \lim_{k\to\infty}\|\mathbf{\Sigma}_{k-1} - \mathcal{S}_{\boldsymbol{w}/\rho_{k-1}}(\mathbf{\Sigma}_{k-1})\|_F + \|\mathbf{X}_{k+1} - \mathbf{X}_k\|_F + \rho_k^{-1}\|\mathbf{A}_{k-1} + \mathbf{A}_{k+1} - \mathbf{A}_k\|_F = 0$, where $\mathbf{U}_{k-1}\mathbf{\Sigma}_{k-1}\mathbf{V}_{k-1}^\top$ is the SVD of the matrix $\mathbf{X}_k + \rho_{k-1}\mathbf{A}_{k-1}$. □

References

[1] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006. 1, 2

[2] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. *IEEE International Conference on Computer Vision (ICCV)*, pages 2272–2279, 2009. 1, 2

[3] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4):1620–1630, 2013. 1

[4] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 60–65, 2005. 1, 2

[5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007. 1, 2, 5

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space. *IEEE*

ICCV
#572

ICCV
#572

ICCV 2017 Submission #572. CONFIDENTIAL REVIEW COPY. DO NOT DISTRIBUTE.

International Conference on Image Processing (ICIP), pages 313–316, 2007. 1, 2, 5

[7] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng. Patch group based nonlocal self-similarity prior learning for image denoising. *IEEE International Conference on Computer Vision (ICCV)*, pages 244–252, 2015. 1, 2

[8] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2862–2869, 2014. 1, 2, 4, 5

[9] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising: Can plain neural networks compete with BM3D? *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2392–2399, 2012. 1, 2, 5

[10] U. Schmidt and S. Roth. Shrinkage fields for effective image restoration. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2774–2781, June 2014. 1, 2, 5

[11] Y. Chen, W. Yu, and T. Pock. On learning optimized reaction diffusion processes for effective image restoration. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5261–5269, 2015. 1, 2, 5

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

[13] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. *IEEE International Conference on Computer Vision (ICCV)*, pages 479–486, 2011. 1, 2

[14] Julien Mairal, Michael Elad, and Guillermo Sapiro. Sparse representation for color image restoration. *IEEE Transactions on image processing*, 17(1):53–69, 2008. 1, 2

[15] Hakki Can Karaimer and Michael S. Brown. A software platform for manipulating the camera imaging pipeline. *European Conference on Computer Vision (ECCV)*, October 2016. 1

[16] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, January 2011. 1, 2

[17] Canyi Lu, Changbo Zhu, Chunyan Xu, Shuicheng Yan, and Zhouchen Lin. Generalized singular value thresholding. *AAAI*, 2015. 1

[18] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Review*, 52(3):471–501, 2010. 2

[19] Maryam Fazel. Matrix rank minimization with applications. *PhD thesis, Stanford University*, 2002. 2

[20] Nathan Srebro, Tommi Jaakkola, et al. Weighted low-rank approximations. *ICML*, 3(2003):720–727, 2003. 2

[21] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010. 2

[22] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011. 2

[23] Zhouchen Lin, Risheng Liu, and Zhixun Su. Linearized alternating direction method with adaptive penalty for low-rank representation. *Advances in neural information processing systems*, pages 612–620, 2011. 2

[24] Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936. 2

[25] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(9):2117–2130, Sept 2013. 2

[26] T. H. Oh, Y. W. Tai, J. C. Bazin, H. Kim, and I. S. Kweon. Partial sum minimization of singular values in robust pca: Algorithm and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(4):744–758, April 2016. 2

[27] Shuhang Gu, Qi Xie, Deyu Meng, Wangmeng Zuo, Xiangchu Feng, and Lei Zhang. Weighted nuclear norm minimization and its applications to low level vision. *International Journal of Computer Vision*, pages 1–26, 2016. 2, 3, 4, 7

[28] S. Nam, Y. Hwang, Y. Matsushita, and S. J. Kim. A holistic approach to cross-channel image noise modeling and its application to image denoising. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1683–1691, 2016. 2, 5, 6, 7

[29] 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. 2

[30] Z. Gong, Z. Shen, and K.-C. Toh. Image restoration with mixed or unknown noises. *Multiscale Modeling & Simulation*, 12(2):458–487, 2014. 2

[31] F. Zhu, G. Chen, and P.-A. Heng. From noise modeling to blind image denoising. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 2

[32] M. Lebrun, M. Colom, and J.-M. Morel. Multiscale image blind denoising. *IEEE Transactions on Image Processing*, 24(10):3149–3161, 2015. 2, 5

[33] M. Lebrun, M. Colom, and J. M. Morel. The noise clinic: a blind image denoising algorithm. http://www.ipol.im/pub/art/2015/125/. Accessed 01 28, 2015. 2, 5

[34] Neatlab ABSoft. Neat Image. https://ni.neatvideo.com/home. 2, 5