Image Denoising via Bandwise Adaptive Modeling and Regularization Exploiting Nonlocal Similarity

Ruiqin Xiong, *Member, IEEE*, Hangfan Liu, Xinfeng Zhang, *Member, IEEE*, Jian Zhang, *Member, IEEE*, Siwei Ma, *Member, IEEE*, Feng Wu, *Fellow, IEEE*, and Wen Gao, *Fellow, IEEE*

***Abstract*— This paper proposes a new image denoising algorithm based on adaptive signal modeling and regularization. It improves the quality of images by regularizing each image patch using bandwise distribution modeling in transform domain. Instead of using a global model for all the patches in an image, it employs content-dependent adaptive models to address the non-stationarity of image signals and also the diversity among different transform bands. The distribution model is adaptively estimated for each patch individually. It varies from one patch location to another and also varies for different bands. In particular, we consider the estimated distribution to have non-zero expectation. To estimate the expectation and variance parameters for every band of a particular patch, we exploit the nonlocal correlation in image to collect a set of highly similar patches as the data samples to form the distribution. Irrelevant patches are excluded so that such adaptively learned model is more accurate than a global one. The image is ultimately restored via bandwise adaptive soft-thresholding, based on a Laplacian approximation of the distribution of similar-patch group transform coefficients. Experimental results demonstrate that the proposed scheme outperforms several state-of-the-art denoising methods in both the objective and the perceptual qualities.**

***Index Terms*— Image denoising, transform domain modeling, bandwise modeling, adaptive regularization, adaptive soft-thresholding, nonlocal similarity.**

I. Introduction

With the development of imaging, computing and communication technologies, there has been a rapid growth in image and video applications in recent years. At the same time, there is an increasing demand for high image qualities. However, the signal captured by camera is susceptible to noise during the acquisition and transmission process. Therefore, denoising remains to be an important problem in many image processing tasks and has attracted much research interest in the past decades [1]–[17].

In image denoising, we generally assume the original scene $\mathbf{x}$ is degraded by some additive noise in the acquisition system, as formulated by

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{n}$ is the noise and $\mathbf{y}$ is the captured signal. The goal is to restore the original signal $\mathbf{x}$ from its corrupted version $\mathbf{y}$, with the best quality we can possibly achieve.

Image denoising is a typical ill-posed inverse problem. In order to obtain an estimate that is non-trivially better than $\mathbf{y}$, it is critical to exploit the prior knowledge we know about the original image so that we can regularize the solution and pick up a particular $\mathbf{x}$ that is more suitable than the others. Many image prior models have been developed in literature to characterize the statistical feature of natural images. Early regularization approaches mainly consider the local correlation among pixels, e.g. the spatial continuity and smoothness of image signals. A well-known method of this kind is total variation (TV) regularization [18], [19]. It formulates the statistical fact that image gradient is close to zero at most regions, and therefore it can be regarded as a kind of sparsity in the image gradient domain.

Some other regularization approaches consider the fact that most image signals can be sparsely represented using decorrelation transforms, e.g. the discrete cosine transform (DCT) or the discrete wavelet transform (DWT), so that the signal can be well separated from the noise. Typical schemes of this category include [20]–[25]. Yu and Sapiro [20] divide image into small blocks and apply DCT transform and a local thresholding on the coefficients. Donoho [21] decomposes image into some wavelet subbands and then applies soft-thresholding to the coefficients to reduce noise. Portilla *et al.* [22] propose to model the wavelet coefficients via mixture of Gaussians. Chang *et al.* [23] propose to employ adaptive wavelet thresholding based on data-driven estimation of the distribution parameters of each wavelet subband. Besides the distribution of a single coefficient itself, some works also consider inter-coefficient correlation. Sendur and Selesnick [24] proposes several bivariate shrinkage functions

Manuscript received September 20, 2015; revised April 9, 2016, July 14, 2016, and September 21, 2016; accepted September 21, 2016. Date of publication September 27, 2016; date of current version October 18, 2016. This work was supported in part by the National Basic Research Program of China under Grant 2015CB351800, in part by the National Natural Science Foundation of China under Grant 61370114, Grant 61425026, and Grant 61421062, in part by the China Postdoctoral Science Foundation under Grant 2016T90017, and in part by the Cooperative Medianet Innovation Center. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alin M. Achim.

R. Xiong, H. Liu, J. Zhang, S. Ma, and W. Gao are with the Institute of Digital Media, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China, and also with the National Engineering Laboratory for Video Technology, Peking University, Beijing 100871, China (e-mail: rqxiong@pku.edu.cn; liuhf@pku.edu.cn; jian.zhang@pku.edu.cn; swma@pku.edu.cn; wgao@pku.edu.cn).

X. Zhang is with the Rapid-Rich Object Search Laboratory, Nanyang Technological University, Singapore 637553 (e-mail: xfzhang@ntu.edu.sg).

F. Wu is with the University of Science and Technology of China, Hefei 230026, China (e-mail: fengwu@ustc.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2614160

to address inter-coefficient dependency. Luisier *et al.* [25] introduce an inter-scale orthonormal wavelet thresholding method.

Besides the methods that use fixed transform bases, some more-advanced methods advocate to use transforms that are adaptively learned from the image content [26]–[29]. For example, Muresan and Parks [26] propose to utilize principal component analysis (PCA) for signal decorrelation. Elad and Aharon [27] use over-complete dictionary to represent image patches and propose a K-SVD algorithm to learn optimal dictionary based on a set of training patches so that the sparsity of signal representation can be maximized.

Another well-known local denoising method is the bilateral filtering [30]. It reduces noise by averaging a group of local pixels according to the distances and the intensity similarity to the current pixel. Non-local means (NLM) [31] can be seen as a patch-based generalization of the bilateral filtering. It extends the processing from a locality to the whole image, so as to exploit the repetitive patterns in images. The potential of non-local mean was further demonstrated by [32] and [33].

Since the invention of NLM, extensive research efforts [34]–[52] have been devoted to exploit the non-local similarity for image restoration. These non-local methods have achieved superior performance, compared with the previous local regularization approaches. Among them, BM3D [34] is a well-recognized benchmark. It utilizes non-local block matching and 3-D transform for signal separation and uses coefficient thresholding to reduce the noise. To be concrete, it stacks a set of similar patches of a reference block into a three-dimensional (3D) data cubic, on which a 3D transform is applied and a hard thresholding (in the first stage) or a Wiener filtering (in the second stage) is subsequently performed. Similar to BM3D, Zhang *et al.* [53] proposed a two-stage denoising scheme (called LPG-PCA), but adopted PCA instead of DCT for patch decorrelation. The gathering of non-local similar patches via block matching is also called structural clustering in some works.

Combining the ideas of structural clustering and dictionary learning, K-LLD [35] advocates that similar patches should share similar sub-dictionaries and utilizes the sub-dictionaries for image modeling. LSSC [54] improves the performance of K-SVD via nonlocal sparse model. CSR [55] attempts to unify the local and nonlocal sparsity constraints. Dictionary learning is large-scale and highly non-convex, and the coherence of dictionaries results in instability and imprecise estimation in the nonlinear sparse inverse problem [56]. To overcome such issues, some works apply Gaussian mixture model (GMM) to image patches since GMM leads to a combination of linear estimations [57]–[59]. Chen *et al.* propose PCLR [60] to jointly exploit external prior and internal prior of image patches, using the GMMs learned from clean images to guide patch clustering of the input noisy images.

In a recent line of research, some methods utilize low-rank matrix approximation for image restoration. These methods are generally based on similar structure grouping in nature. Wang *et al.* [61] propose that a matrix consists of non-local similar patches should be of low-rank and have sparse singular values. Ji *et al.* [62] employ nuclear norm minimization (NNM) for video denoising. Gu *et al.* [40] improve the NNM approach by regularizing the singular values with different weights.

This paper proposes to denoise natural images via an adaptive signal modeling and regularization framework. The method regularizes every image patch using bandwise distribution modeling in transform domain. Instead of using a global model, the scheme employs content adaptive models that are estimated for each patch individually. In particular, we consider the possibility of non-zero expectation for the distribution of transform coefficient. To estimate the expectation and variance for each coefficient, we exploit non-local correlation and collect a set of similar patches as data samples to form the distribution. We use PCA to determine the decorrelation transform, and treat different transform bands unequally according to their statistical characteristics. To study the distribution of data in non-local similar-patch group, we considered generalized Gaussian distributions (GGD) and observed that the coefficient distribution of non-local similar patches can be approximated by Laplacian distribution. Bandwise adaptive soft-thresholding (BAS) is ultimately conducted to regularize the coefficient of each band, therefore we name the proposed scheme as BAS denoising.

The remainder of this paper is organized as follows. Section II briefly reviews both the local and nonlocal image denoising methods. Section III explains the bandwise adaptive modeling and regularization framework. Section IV describes the details of the proposed BAS algorithm, including parameter estimation and numerical solution of the optimization problem. Experimental results are reported in Section V and Section VI concludes the paper.

II. Review of Image Denoising Background

In Bayesian framework, optimum solution to the problem (1) is the *maximum a posterior probability* (MAP) estimator

$$\tilde{\mathbf{x}} = \arg\max_{\mathbf{x}} \Pr(\mathbf{x}|\mathbf{y}). \tag{2}$$

According to the Bayes rule, this can be transformed to

$$\begin{aligned}\tilde{\mathbf{x}} &= \arg\max_{\mathbf{x}} \frac{\Pr(\mathbf{y}|\mathbf{x}) \cdot \Pr(\mathbf{x})}{\Pr(\mathbf{y})} \\ &= \arg\max_{\mathbf{x}} \log \Pr(\mathbf{y}|\mathbf{x}) + \log \Pr(\mathbf{x}).\end{aligned} \tag{3}$$

The term $\log \Pr(\mathbf{y}|\mathbf{x})$ indicates the likelihood of $\mathbf{x}$, and $\log \Pr(\mathbf{x})$ reflects the prior knowledge we know about $\mathbf{x}$. In the case of additive white Gaussian noise (AWGN), the MAP estimation problem can be generally formulated as

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \cdot \Psi(\mathbf{x}). \tag{4}$$

Interpreted from the regularization point of view, $\|\mathbf{y} - \mathbf{x}\|_2^2$ is the data fidelity term indicating how an estimate $\mathbf{x}$ conforms to the observation $\mathbf{y}$, and $\Psi(\mathbf{x})$ is a function to regularize the solution so that we may pick up a high quality estimate with certain feature that we prefer. λ is the regularization parameter to control the trade-off between the data fidelity and the regularity we desire.

The topic of choosing a proper $\Psi(\mathbf{x})$ has been at the foundation of image processing research since its early days and there has been an evolution of choices for $\Psi(\mathbf{x})$ through the years [63]. It started with the very simple energy term $\Psi(\mathbf{x}) = \|\mathbf{x}\|_2^2$ and then progressed to the smoothness term $\Psi(\mathbf{x}) = \|\mathbf{L}\mathbf{x}\|_2^2$, with $\mathbf{L}$ being the Laplacian operator to extract some kinds of derivatives or high-frequency components of low energy. A further step was to adaptively demand this low energy with $\Psi(\mathbf{x}) = \|\mathbf{L}\mathbf{x}\|_W^2$, allowing the pixels at edges to contain high energy in derivative, where W is a diagonal matrix to assign different weights to different regions. Another major step was the abandonment of the ℓ_2-norm and adoption of robust statistics [63], generally formulated as $\Psi(\mathbf{x}) = \rho(\mathbf{L}\mathbf{x})$. One example is total variation prior, $\Psi(\mathbf{x}) = \|\nabla\mathbf{x}\|_1$, where ∇ is the gradient operator. Another example is the wavelet sparsity prior $\Psi(\mathbf{x}) = \|\mathbf{W}\mathbf{x}\|_1$, where $\mathbf{W}$ is a wavelet transform.

Many recent image restoration methods are built upon the observation that image signals can be sparsely represented in some transform domains. They generally employ a regularization function of the form

$$\Psi(\mathbf{x}) = \|\Phi\mathbf{x}\|_p^q, \tag{5}$$

where Φ is a transform matrix that can decorrelate the signal. Typical examples of the $\|\cdot\|_p^q$-norm include $\|\cdot\|_2^2$, $\|\cdot\|_1$ and $\|\cdot\|_0$. The ℓ_2-norm is essentially a Gaussian model for the transform coefficients, in which case the problem (4) can be solved by wiener filtering. The ℓ_1-norm corresponds to a Laplacian model for the coefficients, in which case the problem (4) can be solved by soft-thresholding. The solution to the ℓ_0-norm is hard-thresholding.

In recent approaches, the modeling of image signals is usually conducted on the patch level instead of the whole image, because the content in an image can be quite complicated, nonstationary, and hard to describe as a whole using an explicit and relatively simple formulation. For this reason, we usually divide the image into a set of overlapping patches, and model each image patch via a prior or a regularization function. In this case, the $\Psi(\mathbf{x})$ can be expressed as

$$\Psi(\mathbf{x}) = \sum_i \|\Phi_i\mathbf{x}_i\|_p^q. \tag{6}$$

where $\mathbf{x}_i$ is a vectorized patch at location i, and Φ_i is the transform matrix for the i^{th} patch $\mathbf{x}_i$, which may be adaptively chosen for each patch.

III. Regularization via Bandwise Adaptive Modeling

This paper introduces a new denoising method based on bandwise adaptive modeling and regularization. The key ideas of the proposed framework include: 1) The regularization is performed based on distribution modeling in transform domain. 2) Instead of using a global model that assumes the same distribution for all transform coefficients, we employ adaptive models that are estimated for each patch individually and can vary for different locations and different transform bands. 3) Both the expectation and the variance parameters of the distribution are adaptively estimated. In particular, we consider the possibility of *non-zero* expectation. 4) To model the distribution for a particular patch, we use a set of non-locally searched similar patches as the data samples to form the distribution. Dissimilar patches are excluded so that the learned model is not affected by those irrelevant image contents. 5) The adaptive regularization is ultimately implemented by adaptive soft-thresholding. Overlapped patches are aggregated and averaged to generate the final estimated image. The ideas are explained in the following subsections.

A. Utilization of Decorrelation Transform

To pursue the best possible denoising performance, it is crucial to exploit the correlation within the original image signal. Taking advantage of the local correlation, a decorrelation transforms is able to compact most of the signal energy into only a few coefficients while making the other coefficients almost empty, i.e. its value being close to zero. On the contrary, the energy of a white noise will remain equally distributed among its elements as long as the transform is orthogonal. In this way, the image signal can be better separated from the noise in the transform domain. Since the signal in most coefficients are almost empty after the transformation, it becomes feasible to remove the noise thoroughly from these coefficients. Higher diversity in the signal energy distribution will lead to better denoising performance, and the energy diversity of the transform coefficients depends on the strength of correlation within the signal. For decorrelation transform, this paper adopts PCA, which is known for its virtue of being signal adaptive and optimal for energy compaction. The advantage of utilizing decorrelation transform has been analyzed in [64].

B. Bandwise Adaptive Regularization

The most straightforward and simplest transform-domain regularization approach is to minimize the total energy of all coefficients, as expressed by

$$\Psi(\mathbf{x}) = \sum_i \|\Phi_i\mathbf{x}_i\|_2^2, \tag{7}$$

This method deals with all transform bands equally, enforcing the same strength of regularization on the different coefficients of each patch. In this paper, we call this method *uniform* denoising. The regularization parameter in this formulation is independent of the location i, meaning that all the patches in the image share the same distribution model, we call it the global model.

A more efficient way is to process each coefficient *unequally*. This is because statistical characteristics of the coefficients in different bands may vary dramatically. For instance, the variance of coefficients in a low frequency band is usually much more significant than that in a high frequency band. Taking such statistical difference between coefficients into consideration, we come up with *band adaptive* operation that adaptively enforces an appropriate regularization strength on different coefficients. The idea can be formulated by

extending (7) with the variance information:

$$\Psi(\mathbf{x}) = \sum_i \|\Phi_i \mathbf{x}_i\|^2_{W_i} \triangleq \sum_i (\Phi_i \mathbf{x}_i)^{\mathrm{T}} W_i (\Phi_i \mathbf{x}_i), \quad (8)$$

where the diagonal matrix $W_i = \mathrm{diag}\left[\omega_{i,1}, \omega_{i,2}, \ldots, \omega_{i,S}\right]$ consists of the parameters to control the regularization strength for different coefficients of patch $\mathbf{x}_i$, and $S = s \times s$ is the size of patch $\mathbf{x}_i$. The subscripts i of W_i (or $\{\omega_{i,j}\}$) imply that the regularization parameter varies for different patches as well as for different bands.

Interpreted in the Bayesian framework, the regularization parameter matrix W_i should be determined by the variance (i.e. uncertainty) of each transform coefficient in $\mathbf{x}_i$. Following the mathematical derivation in [65], the regularization parameters in Eq. (4) and Eq. (8) should be

$$\lambda = 2\sigma_n^2, \quad \omega_{i,j} = \frac{1}{2\sigma_{i,j}^2}, \quad (9)$$

where σ_n^2 is the variance of noise, whereas $\sigma_{i,j}^2$ is the variance of the j-th coefficient in the transformed patch $\Phi_i \mathbf{x}_i$. Band adaptive regularization in the transform domain is expected to achieve significantly better performance than the band-uniform regularization method. This will be verified by experiments in the Section V-A.

C. Adaptive Expectation Modeling

The regularization approach formulated in Eq. (8) is actually assuming a *zero-mean* Gaussian distribution (with an adaptively estimated variance) as the statistical model for the transform coefficients of patches at all locations. From another point of view, we are using zero as a prediction for the coefficients. However, this model is not good enough. The zero-mean assumption is indeed valid as a *global model* for all the patches *as a whole*, since most regions of an image are locally smooth. However, for an *individual* patch, this global model can be misleading. For example, for regions with rich textures or edges, the middle-frequency and high-frequency coefficients may have significant values so that this zero-mean distribution model is no longer suitable.

Therefore, in addition to exploiting the statistic of variance, we further extend the regularization model (8) to incorporate a more accurate statistic of coefficient expectation, as formulated by

$$\Psi(\mathbf{x}) = \sum_i \left\|\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i\right\|^2_{W_i} \triangleq \sum_i (\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i)^{\mathrm{T}} W_i (\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i) \quad (10)$$

or equivalently

$$\Psi(\mathbf{x}) = \sum_i \sum_j \frac{\left(\Phi_{i,j} \mathbf{x}_i - \mu_{i,j}\right)^2}{2\sigma_{i,j}^2}, \quad (11)$$

where $\boldsymbol{\mu}_i = [\mu_{i,1}, \mu_{i,2}, \ldots, \mu_{i,S}]^{\mathrm{T}}$, $\mu_{i,j}$ is the expectation of the coefficient $\Phi_{i,j} \mathbf{x}_i$, and $\Phi_{i,j}$ is the j-th row of Φ_i. The incorporation of the term $\boldsymbol{\mu}_i$ would effectively reduce the mean square prediction error of the coefficients and thus improve the denoising performance.

The subscript i in Eq. (10) indicates that the regularization is location adaptive. To be specific, the transform matrix Φ_i should be chosen adaptively based on the covariance characteristics of the signal content at each patch location. The matrix W_i and the vector $\boldsymbol{\mu}_i$ should be determined based on the variance and expectation statistics at each patch location. In this way, this framework addresses the non-stationarity of image signals.

D. Distribution Modeling of Coefficients

In order to establish a content adaptive model for a particular patch, say $\mathbf{x}_i$, a plausible way is to collect a set of patches from the image that are similar to this patch and regard them as the data sample of the unknown random vector $\mathbf{x}_i$. Irrelevant patches should be excluded so as to guarantee the modeling accuracy.

To find a proper distribution model for the coefficients in a similar-patch group, we fit the empirical distribution by a family of generalized Gaussian distribution (GGD), which is commonly used in the field of image coding and processing. The generalized Gaussian density has the following form:

$$\mathcal{G}(u; \beta) = \frac{1}{2\Gamma(1 + \frac{1}{\beta})} \exp\left(-|u|^{\beta}\right), \quad (12)$$

where $\Gamma(\cdot)$ is the gamma function and β is the shape parameter which controls the overall shape of the distribution. Particularly, the $\mathcal{G}(u; \beta)$ is a Laplacian distribution when $\beta = 1$ and is a Gaussian distribution when $\beta = 2$.

To study the actual distribution of image data in similar-patch groups, we have extracted more than 6.5×10^4 similar-patch groups from 9 natural images, with every group consisting of 60 similar patches. We calculate the covariance matrix of the patch vectors in each group and apply the corresponding PCA transform to decorrelate each group. Suppose $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$ is the pixel matrix for a particular similar-patch group, where $\mathbf{x}_1$ is a reference patch and other $\{\mathbf{x}_k\}$ are similar patches; Φ is the PCA transform matrix produced by the covariance matrix of X. $C = \Phi X$ is the transform coefficient matrix of this similar-patch group, $C^{(k)}$ (the k-th row of C) is the k-th coefficient band, $\mu^{(k)}$ and $\sigma^{(k)}$ are the mean value and standard deviation of that band. Then we centralize and normalize the coefficient band $C^{(k)}$ by

$$\tilde{C}^{(k)} = \left(C^{(k)} - \mu^{(k)}\right) / \sigma^{(k)}. \quad (13)$$

After that, the centralized and variance-normalized coefficient band $\tilde{C}^{(k)}$ from all similar-patch groups are gathered together as samples to form an overall distribution. This is shown in Fig. 1. We can see that the centralized and variance-normalized coefficients have very similar distribution for all bands.

To study this distribution, we have used generalized Gaussian distribution to analyze it. To find a proper shape parameter β of GGD that can best fit the empirical distribution of PCA transform coefficients, we employ Kullback-Leibler (KL) divergence to measure the conformance, with β ranging from

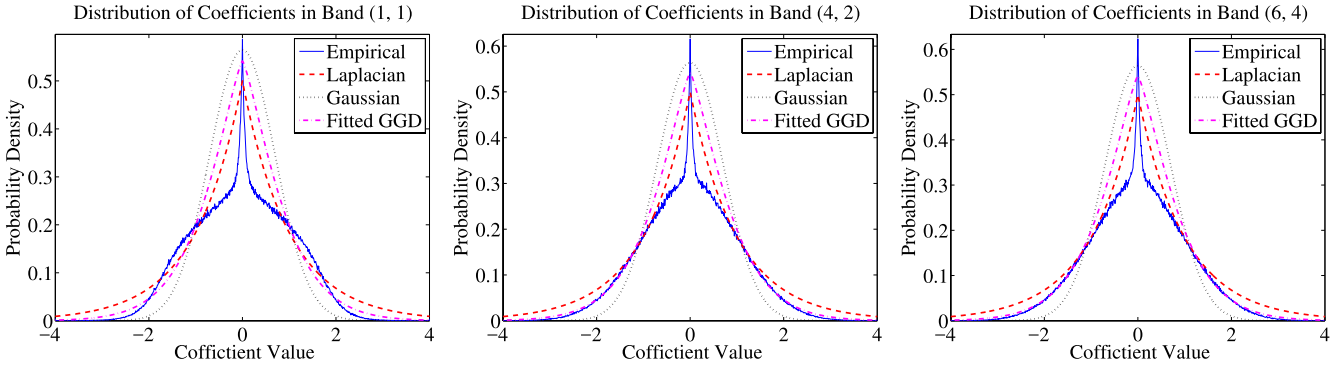


Fig. 1. The centralized and variance-normalized distribution of transform coefficient in similar-patch groups. The empirical distribution is approximated by generalized Gaussian distribution (GGD). A 6×6 transform is used in this experiment and band (u, v) is the u-th and the v-th band in the vertical and the horizontal directions, respectively.

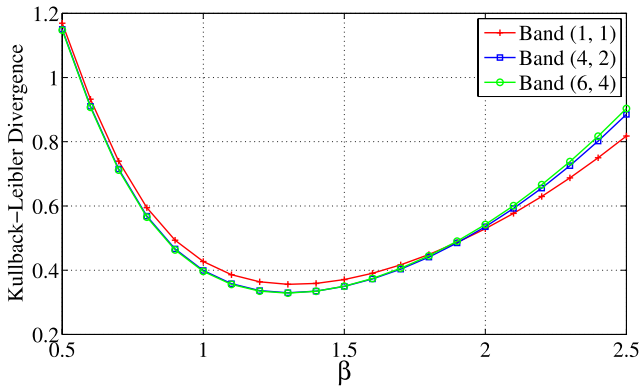


Fig. 2. The Kullback-Leibler divergence between the empirical distribution and its fitted GGD distribution with shape parameter β ranging from 0.5 to 2.5. The three curves correspond to three different bands. The optimal β is within the interval [1, 1.5].

TABLE I
PARAMETER SETTINGS

Noise level	ρ	M	s
$\sigma_n \leq 20$	0.12	80	7
$20 < \sigma_n \leq 40$	0.12	100	8
$40 < \sigma_n \leq 80$	0.12	135	9
$\sigma_n > 80$	0.12	150	10

Fig. 3. The test images used in the paper.

Algorithm 1 The BAS Algorithm

Data: the noisy image $\mathbf{y}$
Result: the denoised image $\tilde{\mathbf{x}}$
Initialization: $\tilde{\mathbf{x}} = \mathbf{y}$, $\tilde{\mathbf{y}} = \mathbf{y}$
while *stop criteria unsatisfied* **do**
  Update $\tilde{\mathbf{y}}$ according to Eq. (30);
  Update the noise variance σ_n^2;
  for $i = 1, 2, 3, \ldots K$ **do**
    Calculate the distances of patches by Eq. (25);
    Group similar patches of $\tilde{\mathbf{y}}_i$ into Y_i;
    Calculate the covariance and PCA matrix for $\mathbf{x}_i$;
    Apply PCA to the patches in Y_i;
    Calculate $\mu_{i,j}$ and $\sigma_{i,j}$;
    Perform the soft-thresholding operation Eq. (21);
    Calculate $\tilde{\mathbf{x}}_i$ by Eq. (23);
  end
  Aggregate the estimated patches to construct $\tilde{\mathbf{x}}$ by Eq. (24);
end

0.5 to 2.5 for every band, as displayed in Fig. 2. We can see that the optimal β falls in the interval [1, 1.5] for all the bands.

We plot the GGD with the best matched β, the matched Laplacian distribution (i.e. $\beta = 1$) as well as the matched Gaussian distribution (i.e. $\beta = 2$) against the empirical distribution in Fig. 1. We can observe that Laplacian distribution can be a good approximation of the empirical distribution. For the convenience of numerical optimization, this paper adopt the Laplacian approximation. The utilization of Laplacian prior has also been suggested in previous works like [59], [61], and [66]. In this way, the regularization function evolves to

$$\begin{aligned}\Psi(\mathbf{x}) &= \sum_i \left\| \Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i \right\|_{1, W_i} \\ &\triangleq \sum_i \left\| W_i \left(\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i \right) \right\|_1, \end{aligned} \tag{14}$$

where $W_i = \operatorname{diag}\left[\omega_{i,1}, \omega_{i,2}, \ldots, \omega_{i,S}\right]$ and $\omega_{i,j} = \sqrt{2}/\sigma_{i,j}$ for $j = 1, 2, \ldots, S$.

IV. THE BAS DENOISING SCHEME

A. Overall Framework

The overall framework of the proposed denoising scheme is grounded on the discussion in Section III. We adopt a patch-based approach, where each patch is first denoised separately and then the estimate of all patches are aggregated to form the final estimate of the whole image. In the denoising stage, each patch is adaptively regularized in transform domain.

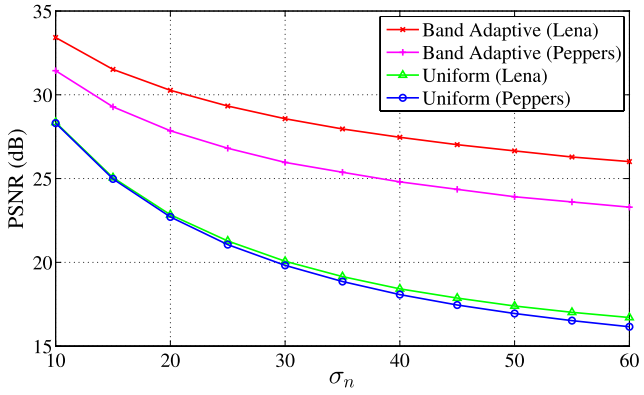


Fig. 4. PSNR comparison between the uniform denoising and the band adaptive denoising. The standard deviation of noise, σ_n, ranges from 10 to 60.

The regularization is bandwise adaptive and varies from one location to another.

We apply the content-adaptive PCA transform to each patch, and utilize nonlocal similar patches in the image as samples to model the distribution of current-patch coefficients. The objective function of the proposed scheme is

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y}-\mathbf{x}\|_2^2 + \lambda \cdot \sum_i \left\|\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i\right\|_{1,W_i}, \quad (15)$$

where $\mathbf{x}_i$ is the vectorized image patch at location i and Φ_i is the PCA transform for $\mathbf{x}_i$. The W_i and $\boldsymbol{\mu}_i$ are learned from the group of similar patches associated with $\mathbf{x}_i$.

The regularization in (15) is applied to image patches, while the data fidelity term is applied to the entire image, making it difficult to solve the optimization problem (15) directly. To tackle this problem, we consider the fact that the patches are regularly extracted from the image so that the times that a pixel appears in the overlapping patches are roughly equal. Therefore, we have

$$\|\mathbf{y}-\mathbf{x}\|_2^2 \approx \eta \sum_i \|\mathbf{y}_i - \mathbf{x}_i\|_2^2 \quad (16)$$

where η is a constant. Thus Eq. (15) is converted to

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \sum_i \left(\|\mathbf{y}_i - \mathbf{x}_i\|_2^2 + \lambda \cdot \left\|\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i\right\|_{1,W_i}\right) \quad (17)$$

with λ updated accordingly. Thus the whole problem can be divided into a set of patch-level estimation sub-problems that can be conquered separately.

B. Patch Based Solution

We separate the optimization problem (17) into patch-level sub-problems

$$\tilde{\mathbf{x}}_i = \arg\min_{\mathbf{x}_i} \|\mathbf{y}_i - \mathbf{x}_i\|_2^2 + \lambda \cdot \left\|\Phi_i \mathbf{x}_i - \boldsymbol{\mu}_i\right\|_{1,W_i}. \quad (18)$$

This is essentially a ℓ_2-ℓ_1 minimization problem which can be easily solved by very simple component-wise soft-thresholding operations. To see this, let $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ be the transform domain representation of $\mathbf{x}_i$ and $\mathbf{y}_i$ respectively, i.e. $\boldsymbol{\alpha}_i = \Phi_i \mathbf{x}_i$, $\boldsymbol{\beta}_i = \Phi_i \mathbf{y}_i$. Then based on the unitary property of Φ_i, we have

$$\|\mathbf{y}_i - \mathbf{x}_i\|_2^2 = \|\Phi_i \mathbf{y}_i - \Phi_i \mathbf{x}_i\|_2^2 = \left\|\boldsymbol{\beta}_i - \boldsymbol{\alpha}_i\right\|_2^2. \quad (19)$$

In this way the optimization problem (18) becomes

$$\begin{aligned}\tilde{\boldsymbol{\alpha}}_i &= \arg\min_{\boldsymbol{\alpha}_i} \left\|\boldsymbol{\beta}_i - \boldsymbol{\alpha}_i\right\|_2^2 + \lambda \cdot \left\|\boldsymbol{\alpha}_i - \boldsymbol{\mu}_i\right\|_{1,W_i} \\ &= \arg\min_{\boldsymbol{\alpha}_i} \sum_j \left(\beta_{i,j} - \alpha_{i,j}\right)^2 + \lambda \cdot \omega_{i,j} \cdot \left|\alpha_{i,j} - \mu_{i,j}\right|, \quad (20)\end{aligned}$$

This can be solved by a soft-thresholding operation [65]:

$$\tilde{\alpha}_{i,j} = \mu_{i,j} + \text{soft}\left(\beta_{i,j} - \mu_{i,j}, \frac{\lambda \omega_{i,j}}{2}\right) \quad (21)$$

with

$$\text{soft}(g,\tau) \triangleq \text{sign}(g) \cdot \max\left(|g| - \tau, 0\right). \quad (22)$$

Then $\mathbf{x}_i$ can be calculated by

$$\tilde{\mathbf{x}}_i = \Phi_i^{\mathrm{T}} \tilde{\boldsymbol{\alpha}}_i. \quad (23)$$

After obtaining the estimate of all patches, we get the full image $\mathbf{x}$ by putting the patches back to their original locations and averaging the overlapped pixels. Suppose R_i is the matrix to extract $\mathbf{x}_i$ from $\mathbf{x}$ at location i, i.e. $\mathbf{x}_i = R_i \mathbf{x}$, then the least-square solution is [27]:

$$\tilde{\mathbf{x}} = \left(\sum_i R_i^{\mathrm{T}} R_i\right)^{-1} \sum_i \left(R_i^{\mathrm{T}} \tilde{\mathbf{x}}_i\right). \quad (24)$$

In this way, the estimate of full image is constructed.

C. Estimation of Model Parameters

In the proposed scheme, the last thing remains to solve is the estimation of the distribution parameters $\mu_{i,j}$ and $\sigma_{i,j}$ for each patch $\mathbf{x}_i$. To this end, we take advantage of the nonlocal similarity within an image, using the nonlocal similar patches as the data samples to form the distribution model of a particular current patch $\mathbf{x}_i$. In addition, we need to learn the signal-adaptive PCA transform Φ_i, which closely relates with the covariance characteristic of the patch $\mathbf{x}_i$. The estimation of these model parameters are conducted as follows.

For a particular current patch $\mathbf{x}_i$ in $\mathbf{x}$, we have $\mathbf{y}_i$ from the noisy input image $\mathbf{y}$, and we search non-locally within $\mathbf{y}$ for its similar patches. Theoretically, this searching should be carried out on the clean image $\mathbf{x}$ so that we can find really-matched patches, but in practice we can only do this on $\mathbf{y}$ or a preliminarily denoised version of $\mathbf{y}$ instead, because the $\mathbf{x}$ is unavailable. The dissimilarity between two patches $\mathbf{x}_i$ and $\mathbf{x}_k$ is measured by the Euclidean distance:

$$d(i,k) = \frac{\|\mathbf{y}_i - \mathbf{y}_k\|_2^2}{s^2}. \quad (25)$$

Then all the patches that has a distance lower than a certain threshold τ are considered to be "similar patches" of $\mathbf{x}_i$ and stacked into a group $\mathbb{G}_i$:

$$\mathbb{G}_i = \{\mathbf{y}_k | d(i,k) \leqslant \tau\}, \quad (26)$$

TABLE II
PSNR COMPARISON OF BM3D [34], PID [67], WNNM [40], LINC [59], PLR [68], PCLR [60] AND THE PROPOSED BAS. (UNIT: *dB*)

σ_n	20							30						
Schemes	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*
Airplane	32.71	32.83	32.98	32.71	32.60	32.91	**33.02**	31.08	31.18	31.31	31.00	30.78	31.22	**31.32**
Barbara	31.24	30.96	31.66	**31.68**	31.40	31.03	31.61	29.02	28.97	**29.53**	29.42	29.23	28.79	29.50
C.man	30.41	30.65	30.68	30.32	30.05	30.72	**30.86**	28.65	28.82	28.78	28.45	28.14	28.77	**29.00**
Lena	31.51	31.51	31.77	31.76	31.46	31.69	**31.88**	29.45	29.53	29.72	29.69	29.41	29.66	**29.82**
Monarch	30.42	30.68	**31.19**	30.70	30.11	31.13	**31.19**	28.37	28.68	28.96	28.57	28.09	28.92	**29.07**
R.R.Hood	32.25	32.19	32.36	32.19	32.09	32.07	**32.50**	30.73	30.69	30.89	30.60	30.45	30.53	**31.00**
Sailboats	33.11	33.18	33.26	33.25	32.98	33.14	**33.41**	31.21	31.28	31.38	31.28	31.00	31.25	**31.49**
Window	32.97	33.13	33.35	33.19	32.70	33.15	**33.52**	30.81	31.03	31.21	30.97	30.46	30.98	**31.39**
Baboon	25.57	25.61	25.67	25.59	25.70	25.72	**25.75**	23.77	23.87	23.96	23.86	**23.99**	23.90	**23.99**
Couple	30.69	30.64	30.75	30.60	30.36	30.68	**30.90**	28.80	28.74	28.92	28.74	28.42	28.84	**29.01**
F.boat	30.81	30.71	30.93	30.76	30.61	30.90	**30.99**	29.02	28.92	29.15	28.95	28.74	29.15	**29.18**
House	33.78	33.61	34.06	33.85	33.42	33.89	**34.13**	32.09	31.94	**32.59**	32.24	31.60	32.22	32.45
Peppers	31.30	31.47	31.57	31.32	30.97	31.57	**31.78**	29.30	29.53	29.52	29.34	29.02	29.58	**29.77**
straw	27.08	27.30	**27.64**	27.51	27.35	27.45	27.62	24.93	25.21	**25.45**	25.39	25.32	25.26	25.44
Average	30.99	31.03	31.28	31.10	30.84	31.15	**31.37**	29.09	29.17	29.38	29.18	28.90	29.22	**29.46**
σ_n	40							50						
Schemes	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*
Airplane	29.79	30.02	**30.13**	29.83	29.52	29.94	30.08	28.81	29.09	**29.22**	28.92	28.51	29.00	29.16
Barbara	27.14	27.54	27.75	27.75	27.63	27.03	**27.96**	26.24	26.41	26.72	26.43	26.36	25.97	**26.81**
C.man	27.20	27.58	27.54	27.24	26.85	27.57	**27.75**	26.12	26.60	26.44	26.31	25.83	26.58	**26.73**
Lena	27.87	28.17	28.29	28.25	27.97	28.24	**28.43**	26.94	27.10	27.33	27.14	26.81	27.18	**27.41**
Monarch	26.73	27.27	27.43	27.10	26.63	27.44	**27.63**	25.70	26.16	26.24	25.96	25.44	26.25	**26.49**
R.R.Hood	29.62	29.71	29.80	29.55	29.35	29.45	**29.99**	28.93	28.97	29.14	28.75	28.53	28.70	**29.23**
Sailboats	29.81	29.94	30.04	29.90	29.57	29.85	**30.13**	28.91	28.91	29.12	28.84	28.47	28.90	**29.13**
Window	29.21	29.51	29.65	29.40	28.90	29.41	**29.84**	28.14	28.33	28.57	28.19	27.74	28.28	**28.67**
Baboon	22.49	22.76	22.89	22.76	**22.91**	22.79	22.87	21.83	21.93	**22.15**	21.91	22.09	22.05	22.11
Couple	27.43	27.38	27.58	27.40	27.05	27.50	**27.67**	26.43	26.31	**26.66**	26.35	26.00	26.55	26.65
F.boat	27.63	27.64	27.83	27.64	27.38	27.86	**27.87**	26.67	26.66	26.84	26.63	26.34	**26.91**	26.89
House	30.68	30.70	**31.42**	30.99	30.15	30.85	31.19	29.70	29.58	**30.46**	29.93	28.98	29.81	30.15
Peppers	27.70	28.09	28.09	27.91	27.57	28.16	**28.32**	26.66	26.93	26.99	26.76	26.41	27.02	**27.16**
straw	23.18	23.69	**23.90**	23.87	23.85	23.74	**23.90**	22.34	22.49	**22.87**	22.62	22.68	22.73	22.81
Average	27.61	27.86	28.02	27.83	27.52	27.85	**28.12**	26.67	26.82	27.05	26.77	26.44	26.85	**27.10**

and the corresponding locations are recorded in a set $\mathbb{L}_i$:

$$\mathbb{L}_i = \{k | d(i,k) \leqslant \tau\}. \tag{27}$$

Alternatively, we may choose to group a fixed number (e.g. M) of most similar patches into $\mathbb{G}_i$.

Once the set of similar patches is obtained, we can calculate the covariance matrix of $\mathbf{x}_i$ and then train the PCA transform Φ_i. Suppose $\mathbf{p}_i^1, \mathbf{p}_i^2, \ldots \mathbf{p}_i^N$ are the vectorized representation of all the similar patches in $\mathbb{G}_i$, and let matrix $P_i = [\mathbf{p}_i^1, \mathbf{p}_i^2, \ldots \mathbf{p}_i^N]$. We calculate mean of the columns in P_i and subtract it from P_i. The covariance matrix of $\mathbf{x}_i$ is estimated by $P_i P_i^{\mathrm{T}}$. The PCA transform matrix Φ_i is then obtained via eigen value decomposition (EVD).

Using the transform Φ_i, the patches in $\mathbb{G}_i$ is converted into PCA domain. The expectation $\mu_{i,j}$ is estimated by the median value of the j-th coefficient band of group $\mathbb{G}_i$, because we assumed a Laplacian approximation. The advantage of such estimation is to exclude the influence of outliers. Furthermore, since the noise and the image signal are independent, the variance of coefficients in the j-th band is estimated by

$$\tilde{\sigma}_{i,j}^2 = \max\left(\sigma_{i,j}^2 - \sigma_n^2, 0\right), \tag{28}$$

where $\sigma_{i,j}^2$ is the variance calculated based on $\mathbf{y}$. Suppose $\beta_{i,j}^k = \Phi_{i,j}\mathbf{y}_k$, $k \in \mathbb{L}_i$, then $\sigma_{i,j}^2$ is calculated by

$$\sigma_{i,j}^2 \doteq \frac{1}{M}\sum_{k\in\mathbb{L}_i}\left(\beta_{i,j}^k - \mu_{i,j}\right)^2, \tag{29}$$

where M is the number of similar patches in the group $\mathbb{G}_i$.

D. Extension to Iterative Regularization

Although the above described is already a complete denoising algorithm, we further explore the idea of iterative denoising as introduced in previous works [55], [69] to pursue better results.

In every iteration, after obtaining an estimation $\tilde{\mathbf{x}}$, we calculate a new noisy observation

$$\mathbf{y}' := \tilde{\mathbf{x}} + \rho(\mathbf{y} - \tilde{\mathbf{x}}) \tag{30}$$

and use it as the noisy image to be further denoised in the next iteration. To get a better understanding of such calculation, we rewrite Eq. (30) as

$$\mathbf{y}' = \mathbf{y} + (1-\rho)\Delta \tag{31}$$

TABLE III
SSIM COMPARISON OF BM3D [34], PID [67], WNNM [40], LINC [59], PLR [68], PCLR [60] AND THE PROPOSED BAS

σ_n	20							30						
Schemes	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*
Airplane	0.8871	0.8885	0.8862	0.8845	0.8868	**0.8893**	0.8871	0.8591	0.8578	0.8551	0.8525	0.8502	**0.8592**	0.8531
Barbara	0.9098	0.9084	0.9150	**0.9174**	0.9137	0.9049	0.9150	0.8605	0.8659	0.8732	0.8730	0.8688	0.8558	**0.8743**
C.man	0.8736	**0.8813**	0.8778	0.8759	0.8737	0.8805	0.8805	0.8358	0.8403	0.8411	0.8325	0.8279	0.8424	**0.8433**
Lena	0.8969	0.8997	0.9015	0.9041	0.8974	0.8966	**0.9026**	0.8558	0.8603	0.8633	**0.8657**	0.8547	0.8613	0.8627
Monarch	0.9212	0.9245	0.9272	0.9252	0.9118	0.9265	**0.9294**	0.8849	0.8931	0.8951	0.8915	0.8761	0.8969	**0.8984**
R.R.Hood	0.8275	0.8240	0.8289	0.8204	0.8262	0.8222	**0.8317**	0.7836	0.7788	0.7869	0.7728	0.7766	0.7787	**0.7878**
Sailboats	0.8916	0.8946	0.8922	**0.8951**	0.8935	0.8895	0.8926	0.8584	**0.8612**	0.8594	0.8606	0.8580	0.8587	0.8576
Window	0.9188	0.9228	0.9224	0.9237	0.9154	0.9194	**0.9250**	0.8826	0.8903	0.8909	0.8902	0.8766	0.8882	**0.8929**
Baboon	0.7217	0.7416	0.7305	0.7377	**0.7522**	0.7382	0.7379	0.6209	0.6313	0.6387	0.6304	**0.6493**	0.6314	0.6394
Couple	0.8441	0.8380	0.8418	0.8358	0.8334	0.8377	**0.8451**	0.7902	0.7814	**0.7928**	0.7828	0.7733	0.7869	0.7921
F.boat	0.8245	0.8194	0.8251	0.8188	0.8211	0.8233	**0.8259**	0.7766	0.7667	**0.7775**	0.7678	0.7665	0.7764	0.7752
House	0.8744	0.8736	0.8730	0.8687	0.8704	0.8730	**0.8801**	0.8492	0.8494	**0.8527**	0.8480	0.8422	0.8496	0.8518
Peppers	0.8873	0.8910	0.8908	0.8884	0.8817	0.8906	**0.8942**	0.8516	0.8596	0.8560	0.8548	0.8460	0.8606	**0.8620**
straw	0.8973	0.9097	0.9106	**0.9129**	0.9114	0.9053	0.9095	0.8275	0.8491	0.8521	0.8551	**0.8568**	0.8425	0.8498
Average	0.8697	0.8726	0.8731	0.8720	0.8706	0.8712	**0.8755**	0.8240	0.8275	0.8311	0.8270	0.8231	0.8278	**0.8315**
σ_n	40							50						
Schemes	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*	*BM3D*	*PID*	*WNNM*	*LINC*	*PLR*	*PCLR*	*BAS*
Airplane	0.8376	**0.8387**	0.8283	0.8323	0.8247	0.8297	0.8246	0.8226	**0.8236**	0.8158	0.8163	0.8022	0.8109	0.8077
Barbara	0.8024	0.8217	0.8204	0.8249	0.8209	0.7969	**0.8301**	0.7624	0.7760	0.7858	0.7736	0.7698	0.7537	**0.7895**
C.man	0.8055	0.8142	0.8111	0.8056	0.7950	**0.8176**	0.8138	0.7836	0.7950	0.7911	0.7846	0.7683	**0.7971**	0.7927
Lena	0.8165	0.8269	0.8284	**0.8321**	0.8172	0.8286	0.8287	0.7893	0.7969	**0.8060**	0.8028	0.7815	0.8030	0.8013
Monarch	0.8477	0.8629	0.8612	0.8597	0.8412	0.8657	**0.8674**	0.8197	0.8336	0.8354	0.8311	0.8061	0.8396	**0.8401**
R.R.Hood	0.7490	0.7497	0.7521	0.7417	0.7411	0.7449	**0.7547**	0.7271	0.7292	**0.7339**	0.7198	0.7144	0.7242	0.7311
Sailboats	0.8278	**0.8320**	0.8266	0.8301	0.8237	0.8253	0.8240	0.8093	0.8074	**0.8097**	0.8040	0.7923	0.8039	0.7994
Window	0.8475	0.8603	0.8586	0.8592	0.8409	0.8555	**0.8612**	0.8235	0.8331	**0.8382**	0.8311	0.8083	0.8293	0.8354
Baboon	0.5377	0.5438	0.5645	0.5472	**0.5690**	0.5538	0.5621	0.4701	0.4682	**0.5083**	0.4728	0.4983	0.5000	0.5039
Couple	0.7424	0.7322	**0.7471**	0.7362	0.7228	0.7405	0.7460	0.7043	0.6885	**0.7144**	0.6948	0.6779	0.7053	0.7056
F.boat	0.7329	0.7236	0.7350	0.7253	0.7200	**0.7356**	0.7315	0.6991	0.6891	0.7029	0.6900	0.6808	**0.7043**	0.6971
House	0.8270	0.8323	**0.8352**	0.8320	0.8188	0.8327	0.8312	0.8149	0.8151	**0.8248**	0.8170	0.7958	0.8190	0.8161
Peppers	0.8158	0.8310	0.8238	0.8246	0.8114	0.8316	**0.8321**	0.7891	0.8033	0.7996	0.7961	0.7793	**0.8058**	0.8045
straw	0.7436	0.7768	0.7848	0.7854	**0.7929**	0.7730	0.7821	0.6812	0.6964	**0.7310**	0.7035	0.7214	0.7194	0.7205
Average	0.7809	0.7890	0.7912	0.7883	0.7814	0.7880	**0.7921**	0.7497	0.7540	**0.7641**	0.7527	0.7426	0.7582	0.7603

TABLE IV
COMPUTATIONAL TIME COMPARISON OF BM3D [34], PID [67], WNNM [40], LINC [59], PLR [68], PCLR [60] AND THE PROPOSED BAS. UNIT: MINUTE(S)

Image size	BM3D	PID	WNNM	LINC	PLR	PCLR	BAS
256×256	0.06	10.77	17.74	18.18	0.14	5.96	11.34
512×512	0.18	40.79	71.58	66.45	0.73	38.76	45.06
768×512	0.25	56.01	109.66	95.26	0.76	53.83	67.85

with $\Delta = \tilde{\mathbf{x}} - \mathbf{y}$. Δ is the change that is supposed to be applied on $\mathbf{y}$ by the basic denoising scheme of one step. Using Eq. (31), we only apply a small fraction of the change. The motivation is to break the "optimal" estimation into multiple stages of weaker denoising. Another way to interpret (30) is to reformulate it as

$$\mathbf{y}' = (1-\rho)\tilde{\mathbf{x}} + \rho\mathbf{y}, \tag{32}$$

i.e. the new "noisy" image $\mathbf{y}'$ is a combination of the old noisy image $\mathbf{y}$ and the estimate $\tilde{\mathbf{x}}$. That means the estimate is partially used.

In the subsequent iterations, we repeat the procedures of transform coefficient modeling and bandwise adaptive soft-thresholding, working on the patches extracted from the updated noisy image $\mathbf{y}'$, and produce the next version of estimate $\tilde{\mathbf{x}}$ by Eq. (24). It is worth noting the value of noise variance σ_n^2 also changes in each iteration. Such technique has appeared in previous works [55], [69].

Now we have described the complete denoising scheme, as summarized in Algorithm 1.

V. EXPERIMENTAL RESULTS

This section evaluates the effectiveness of the proposed method. We divide this section into two parts. In the first part, we verify the conjecture that the band-adaptive denoising performs better than the so-called band-uniform denoising scheme. In the second part, we compare the proposed approach with several benchmark denoising algorithms.

We evaluate the performance of different denoising schemes via peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [70]. Given a ground truth image $\mathbf{x}$, the PSNR of a denoised image $\tilde{\mathbf{x}}$ is defined by

$$\mathrm{PSNR}(\mathbf{x}, \tilde{\mathbf{x}}) = 10 \cdot \log_{10}\left(\frac{255^2}{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2}\right). \tag{33}$$

Fig. 5. Restoration results for *Red Riding Hood*. From left to right: (a) Noisy image (σ_n = 20); (b) BM3D (PSNR = 32.25dB); (c) PID (PSNR = 32.19dB); (d) WNNM (PSNR = 32.36dB); (e) LINC (PSNR = 32.19dB); (f) PLR (PSNR = 32.09dB); (g) PCLR (PSNR = 32.07dB); (h) The proposed BAS (PSNR = **32.50**dB). Please enlarge the figure for better comparison.

Given an image patch $\mathbf{u}$ from $\mathbf{x}$ and the corresponding patch $\mathbf{v}$ from $\tilde{\mathbf{x}}$, the local SSIM index is calculated by:

$$\mathrm{SSIM}(\mathbf{u},\mathbf{v}) = \frac{(2\mu_u\mu_v + C_1)(2\sigma_{uv} + C_2)}{\left(\mu_u^2 + \mu_v^2 + C_1\right)\left(\sigma_u^2 + \sigma_v^2 + C_2\right)}, \quad (34)$$

where μ_u and σ_u are the mean intensity and standard deviation of $\mathbf{u}$ respectively, σ_{uv} represents the cross correlation between $\mathbf{u}$ and $\mathbf{v}$, and C_1, C_2 are two constants used to avoid instability. The SSIM value of the whole image is obtained by averaging the local SSIM indices using a sliding window [70].

A. Band Adaptive Denoising vs. Band Uniform Denoising

We have explained why band adaptive denoising performs better in Section III. This section experimentally demonstrates the advantage of band adaptive denoising over band-uniform denoising.

In order to make fair comparison for the denoising results of Eq. (7) and Eq. (8), which respectively represent band adaptive approach and uniform denoising, we let both techniques follow exactly the same patch-based denoising procedure except for the calculation and utilization of regularization parameters, which are decided by variance of coefficients as indicated in Eq. (9). For band adaptive denoising, after the patches are transformed, coefficients in different bands are separately gathered and used as data samples to estimate the variance values $\sigma_{i,j}^2$. For band-uniform denoising, all the coefficients, no matter which band they belong to, are gathered to calculate the variance σ^2. As a result, in band-adaptive denoising, the noisy coefficients in different bands are adaptively filtered according to their individual variance; while in band-uniform denoising, all the coefficients are processed with the same filter.

We have tested the denoising performance on *Lena* and *Peppers*, with the standard deviation of noise σ_n ranging from 10 to 60, as shown in Fig. 4. Others images produce similar results. It is evident that band adaptive regularization leads to better denoising performance.

B. Performance Comparison With the State-of-the-Art Methods

This section compares the proposed approach with six recent denoising schemes, including BM3D [34], PID [67], WNNM [40], LINC [59], PLR [68], and PCLR [60]. The proposed algorithm is implemented in MATLAB, and the six anchor schemes are tested using the executables or source codes provided by the authors of these methods. In our implementation, we have used the PID [67] method to produce a preliminary result, which is used for block matching and covariance matrix estimation in the first iteration of our algorithm. Parameters used in the proposed algorithm are empirically chosen according to the noise levels. The setting of three parameters are shown in Table I, including the patch size s^2, the number M of similar patches in a group, and the iterative regularization parameter ρ used in (30). The denoising schemes are tested on 14 typical natural images with a wide range of noise levels. The test images are displayed in Fig 3.

As shown in Table II, the proposed scheme outperforms the other six methods in most cases in terms of PSNR. The average PSNR of the proposed method is about 0.4dB higher than

Fig. 6. Restoration results for *Barbara*. From left to right: (a) Noisy image ($\sigma_n = 40$); (b) BM3D (PSNR = 27.14dB); (c) PID (PSNR = 27.54dB); (d) WNNM (PSNR = 27.75dB); (e) LINC (PSNR = 27.75dB); (f) PLR (PSNR = 27.63dB); (g) PCLR (PSNR = 27.03dB); (h) The proposed BAS (PSNR = **27.96**dB).

Fig. 7. Restoration results for *Window*. From left to right: (a) Noisy image ($\sigma_n = 50$); (b) BM3D (PSNR = 28.14dB); (c) PID (PSNR = 28.33dB); (d) WNNM (PSNR = 28.57dB); (e) LINC (PSNR = 28.19dB); (f) PLR (PSNR = 27.74dB); (g) PCLR (PSNR = 28.28dB); (h) The proposed BAS (PSNR = **28.67**dB). Please enlarge the figure for better comparison.

BM3D, 0.3dB higher than PID, 0.1dB higher than WNNM, 0.3dB higher than LINC, 0.6dB higher than PLR, and 0.25dB higher than PCLR. The SSIM result of BAS is also highly competitive against the other anchors, as displayed in Table III. Besides, we show the computational time of each method in Table IV. Note that the BM3D is implemented in C while the others are implemented in MATLAB. In practice, the computation efficiency may be further improved via implementation optimization or parallel computing techniques.

Since the processed images are to be viewed by human eyes, the goal is to achieve better perceptual quality. Fig. 5, 6, 7 and 8 demonstrate the improvement of the

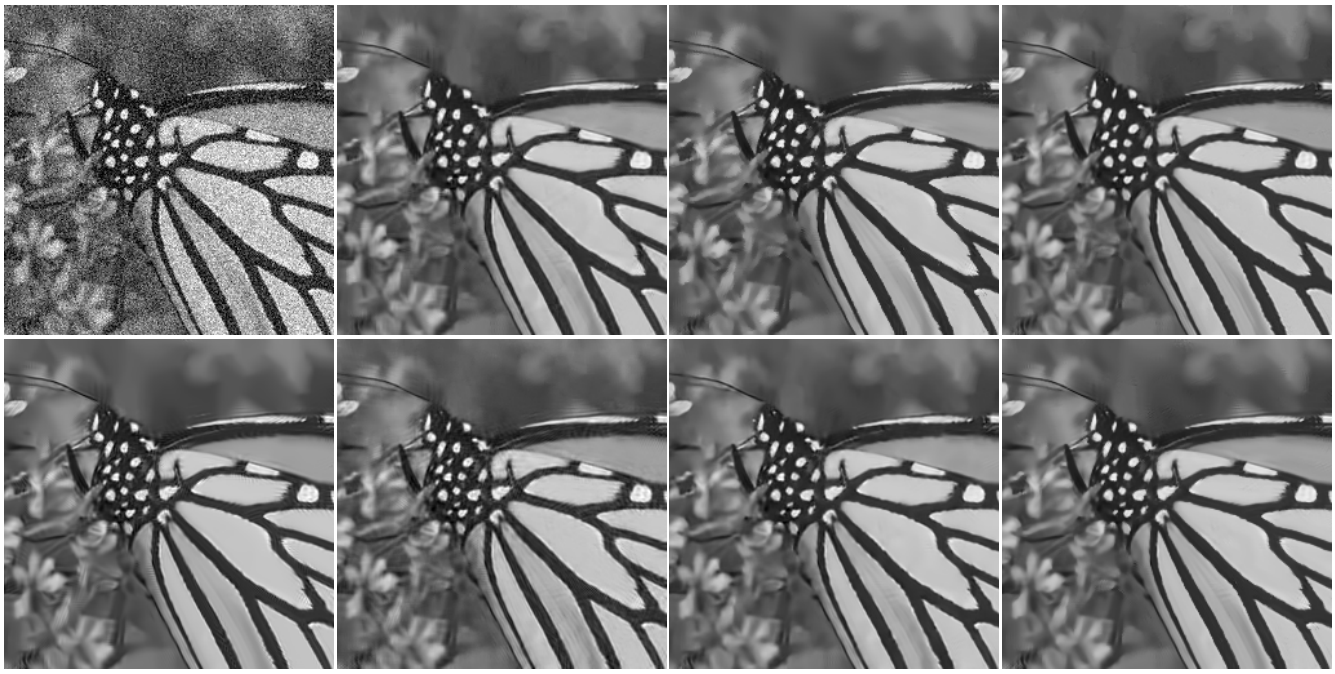


Fig. 8. Restoration results for *Monarch*. From left to right: (a) Noisy image ($\sigma_n = 30$); (b) BM3D (PSNR = 28.37dB); (c) PID (PSNR = 28.68dB); (d) WNNM (PSNR = 28.96dB); (e) LINC (PSNR = 28.57dB); (f) PLR (PSNR = 28.09dB); (g) PCLR (PSNR = 28.92dB); (h) The proposed BAS (PSNR = **29.07**dB). Please enlarge the figure for better comparison.

proposed scheme in visual quality at different noise levels. It is evident that the images produced by the proposed BAS scheme exhibit less noise and artifacts (e.g. in the face area in Fig. 5), and better preserve the details and textures (e.g. the textures on the left side of *Barbara* in Fig. 6, and the pattern on the bottom of *Window* in Fig. 7). Compared with the six anchor schemes, the output of the BAS scheme is visually more pleasant.

VI. Conclusions

This paper develops an new image denoising scheme based on bandwise adaptive regularization. The proposed method adaptively models the actual transform-domain distribution of every patch, using a set of nonlocal similar patches as data samples to form the empirical distribution. Since irrelevant patches are excluded, the proposed model is more accurate than a globally learned model. The regularization is implemented via bandwise adaptive soft-thresholding, based on a Laplacian approximation of the adaptively learned coefficient distribution. Experimental results demonstrate that the proposed denoising scheme outperforms several state-of-the-art methods in terms of both objective quality and perceptual quality.

In this paper, we studied the empirical distribution of the transform coefficients in a similar-patch group, and used a generalized Gaussian distribution (GGD) model to analyze it. We note that the actual distribution is not generalized Gaussian exactly. The reason why we adopt a Laplacian model in the paper is that it is a good approximation with elegant formulation so that the optimization problem can be easily solved. In fact, the actual coefficient distribution is more like a combination of Laplacian (in the middle) and Gaussian (on the two sides). This aspect may be investigated in future works.

Acknowledgment

The authors would like to thank the authors of BM3D [34], WNNM [40], LINC [59], PCLR [60], PID [67] and PLR [68] for providing their source codes or executables for performance comparison.

References

[1] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[2] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman, "Automatic estimation and removal of noise from a single image," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 299–314, Feb. 2008.

[3] M. Zhang and B. K. Gunturk, "Multiresolution bilateral filtering for image denoising," *IEEE Trans. Image Process.*, vol. 17, no. 12, pp. 2324–2333, Dec. 2008.

[4] F. Luisier, T. Blu, and M. Unser, "Image denoising in mixed Poisson–Gaussian noise," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 696–708, Mar. 2011.

[5] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 479–486.

[6] P. Chatterjee and P. Milanfar, "Patch-based near-optimal image denoising," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1635–1649, Apr. 2012.

[7] J. Zhang, R. Xiong, C. Zhao, S. Ma, and D. Zhao, "Exploiting image local and nonlocal consistency for mixed Gaussian-impulse noise removal," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2012, pp. 592–597.

[8] Y. Chen and K. J. R. Liu, "Image denoising games," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 10, pp. 1704–1716, Oct. 2013.

[9] A. Rajwade, A. Rangarajan, and A. Banerjee, "Image denoising using the higher order singular value decomposition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 849–862, Apr. 2013.

[10] M. Mäkitalo and A. Foi, "Noise parameter mismatch in variance stabilization, with an application to Poisson–Gaussian noise estimation," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5348–5359, Dec. 2014.

[11] W. Zuo, L. Zhang, C. Song, D. Zhang, and H. Gao, "Gradient histogram estimation and preservation for texture enhanced image denoising," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2459–2472, Jun. 2014.

[12] H. Peng, R. Rao, and S. A. Dianat, "Multispectral image denoising with optimized vector bilateral filter," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 264–273, Jan. 2014.

[13] H. Talebi and P. Milanfar, "Global image denoising," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 755–768, Feb. 2014.

[14] J. Zhang, D. Zhao, R. Xiong, S. Ma, and W. Gao, "Image restoration using joint statistical modeling in a space-transform domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 6, pp. 915–928, Jun. 2014.

[15] C.-T. Huang, "Bayesian inference for neighborhood filters with application in denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1657–1665.

[16] F. Chen, X. Zeng, and M. Wang, "Image denoising via local and nonlocal circulant similarity," *J. Vis. Commun. Image Represent.*, vol. 30, pp. 117–124, Jul. 2015.

[17] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot, "A convex approach for image restoration with exact Poisson–Gaussian likelihood," *SIAM J. Imag. Sci.*, vol. 8, no. 4, pp. 2662–2682, 2015.

[18] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[19] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[20] G. Yu and G. Sapiro, "DCT image denoising: A simple and effective image denoising algorithm," *Image Process. On Line*, vol. 1, Oct. 2011. [Online]. Available: http://dx.doi.org/10.5201/ipol.2011.ys-dct

[21] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[22] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[23] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.

[24] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.

[25] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 593–606, Mar. 2007.

[26] D. D. Muresan and T. W. Parks, "Adaptive principal components and image denoising," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2003, pp. 101–104.

[27] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[28] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[29] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.

[30] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jan. 1998, pp. 839–846.

[31] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 60–65.

[32] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2866–2878, Oct. 2006.

[33] C. Kervrann and J. Boulanger, "Local adaptivity to variable smoothness for exemplar-based image regularization and representation," *Int. J. Comput. Vis.*, vol. 79, no. 1, pp. 45–69, Aug. 2008.

[34] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[35] P. Chatterjee and P. Milanfar, "Clustering-based denoising with locally learned dictionaries," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1438–1451, Jul. 2009.

[36] J. Zhang, R. Xiong, S. Ma, and D. Zhao, "High-quality image restoration from partial random samples in spatial domain," in *Proc. IEEE Int. Conf. Vis. Commun. Image Process.*, Nov. 2011, pp. 1–4.

[37] M. Lebrun, A. Buades, and J.-M. Morel, "A nonlocal Bayesian image denoising algorithm," *SIAM J. Imag. Sci.*, vol. 6, no. 3, pp. 1665–1688, 2013.

[38] X. Zhang, R. Xiong, X. Fan, S. Ma, and W. Gao, "Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4613–4626, Dec. 2013.

[39] H. Liu, R. Xiong, S. Ma, X. Fan, and W. Gao, "Non-local extension of total variation regularization for image restoration," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2014, pp. 1102–1105.

[40] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2014, pp. 2862–2869.

[41] J. Zhang, D. Zhao, and W. Gao, "Group-based sparse representation for image restoration," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3336–3351, Aug. 2014.

[42] H. Liu, R. Xiong, S. Ma, X. Fan, and W. Gao, "Gradient based image transmission and reconstruction using non-local gradient sparsity regularization," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2014, pp. 1–6.

[43] H. Liu, R. Xiong, S. Ma, X. Fan, and W. Gao, "Gradient based image/video softcast with grouped-patch collaborative reconstruction," in *Proc. IEEE Vis. Commun. Image Process. Conf.*, Dec. 2014, pp. 141–144.

[44] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, and W. Gao, "Image compressive sensing recovery via collaborative sparsity," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 2, no. 3, pp. 380–391, Sep. 2012.

[45] H. Liu, R. Xiong, J. Zhang, and W. Gao, "Image denoising via adaptive soft-thresholding based on non-local samples," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 484–492.

[46] R. Xiong, J. Zhang, F. Wu, and W. Gao, "High quality image reconstruction via non-local collaborative estimation for wireless image/video softcast," in *Proc. Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 2542–2546.

[47] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, and W. Gao, "Compressed sensing recovery via collaborative sparsity," in *Proc. IEEE Data Compress. Conf. (DCC)*, Apr. 2012, pp. 287–296.

[48] X. Zhang, R. Xiong, W. Lin, S. Ma, J. Liu, and W. Gao, "Video compression artifact reduction via spatio-temporal multi-hypothesis prediction," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 6048–6061, Dec. 2015.

[49] J. Zhang, R. Xiong, C. Zhao, Y. Zhang, S. Ma, and W. Gao, "CONCOLOR: Constrained non-convex low-rank model for image deblocking," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1246–1259, Mar. 2016.

[50] H. Liu, R. Xiong, X. Zhang, Y. Zhang, S. Ma, and W. Gao, "Non-local gradient sparsity regularization for image restoration," *IEEE Trans. Circuits Syst. Video Technol.*, to be published. [Online]. Available: http://ieeexplore.ieee.org/document/7457297/, doi: 10.1109/TCSVT.2016.2556498.

[51] X. Zhang, W. Lin, R. Xiong, X. Liu, S. Ma, and W. Gao, "Low-rank decomposition-based restoration of compressed images via adaptive noise estimation," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4158–4171, Sep. 2016.

[52] X. Zhang *et al.*, "Low-rank based nonlocal adaptive loop filter for high efficiency video compression," *IEEE Trans. Circuits Syst. Video Technol.*, to be published. [Online]. Available: http://ieeexplore.ieee.org/document/7492175/, doi: 10.1109/TCSVT.2016.2581618.

[53] L. Zhang, W. Dong, D. Zhang, and G. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping," *Pattern Recognit.*, vol. 43, no. 4, pp. 1531–1549, 2010.

[54] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2272–2279.

[55] W. Dong, X. Li, D. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 457–464.

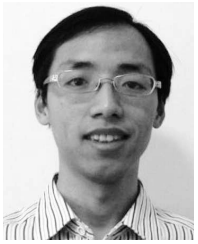
[56] S. Mallat and Y. Guoshen, "Super-resolution with sparse mixing estimators," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2889–2900, Nov. 2010.

[57] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2481–2499, May 2012.

[58] Y.-Q. Wang and J.-M. Morel, "SURE guided Gaussian mixture image denoising," *SIAM J. Imag. Sci.*, vol. 6, no. 2, pp. 999–1034, 2013.

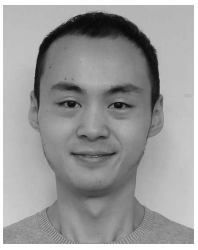
[59] M. Niknejad, H. Rabbani, and M. Babaie-Zadeh, "Image restoration using Gaussian mixture models with spatially constrained patch clustering," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3624–3636, Nov. 2015.

[60] F. Chen, L. Zhang, and H. Yu, "External patch prior guided internal clustering for image denoising," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 603–611.
[61] S. Wang, L. Zhang, and Y. Liang, "Nonlocal spectral prior model for low-level vision," in *Proc. 11th Asian Conf. Comput. Vis.*, Nov. 2012, pp. 231–244.
[62] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1791–1798.
[63] E. Michael, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer, 2010.
[64] R. Xiong, F. Wu, J. Xu, X. Fan, C. Luo, and W. Gao, "Analysis of decorrelation transform gain for uncoded wireless image and video communication," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1820–1833, Apr. 2016.
[65] I. Selesnick. (May 2013). *A Derivation of the Soft-Thresholding Function*. [Online]. Available: http://eeweb.poly.edu/iselesni/lecture_notes/SoftThresholding.pdf
[66] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, "Patch group based nonlocal self-similarity prior learning for image denoising," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 244–252.
[67] C. Knaus and M. Zwicker, "Progressive image denoising," *IEEE Trans. Image Process.*, vol. 23, no. 7, pp. 3114–3125, Jul. 2014.
[68] H. Hu, J. Froment, and Q. Liu. (Jun. 2015). "Patch-based low-rank minimization for image denoising." [Online]. Available: http://arxiv.org/abs/1506.08353
[69] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 460–489, 2005.
[70] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.



Ruiqin Xiong (M'08) received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 2001, and the Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2007.

He was a Research Intern with Microsoft Research Asia from 2002 to 2007, and a Senior Research Associate with the University of New South Wales, Australia, from 2007 to 2009. He joined the Institute of Digital Media, National Engineering Laboratory for Video Technology, Peking University, in 2010, where he is currently a Professor. He has authored over 80 technical papers in referred international journals and conference proceedings. His research interests include statistical image modeling, image and video processing, and video compression and communication.



Hangfan Liu received the B.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 2011. He is currently pursuing the Ph.D. degree with Peking University, Beijing, China. His research interests include image processing and visual communication.

Xinfeng Zhang (M'16) received the B.S. degree in computer science from the Hebei University of Technology, Tianjin, China, in 2007, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2014.

He was a Research Assistant with the Institute of Digital Media, Peking University, from 2009 to 2014. He is currently a Research Fellow with Nanyang Technological University, Singapore. His research interests include image and video processing, and image and video compression.

Jian Zhang (M'14) received the B.Sc. degree from the Department of Mathematics, Harbin Institute of Technology (HIT), Harbin, China, in 2007, and the M.Eng. and Ph.D. degrees from the School of Computer Science and Technology, HIT, in 2009 and 2014, respectively. He is a currently a Post-Doctoral Fellow with the Institute of Digital Media, Peking University, Beijing, China.

His research interests include image/video compression and restoration, compressive sensing, sparse representation, and deep learning. He was a recipient of the Best Paper Award and the Best Student Paper Award at the IEEE International Conference on Visual Communication and Image Processing in 2011 and 2015, respectively.

Siwei Ma (M'12) received the B.S. degree from Shandong Normal University, Jinan, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005.

From 2005 to 2007, he held a post-doctoral position with the University of Southern California. He joined the Institute of Digital Media, EECS, Peking University, where he is currently a Professor. He has authored over 100 technical articles in refereed journals and proceedings in the areas of image and video coding, video processing, video streaming, and transmission.

Feng Wu (M'99–SM'06–F'13) received the B.S. degree in electrical engineering from Xidian University in 1992, and the M.S. and Ph.D. degrees in computer science from the Harbin Institute of Technology in 1996 and 1999, respectively. He was a Principle Researcher and a Research Manager with Microsoft Research Asia. He is currently a Professor with the University of Science and Technology of China.

He holds 77 granted U.S. patents. He has authored or co-authored over 200 high-quality papers, including several dozens of the IEEE transactions papers and top conference papers on MOBICOM, SIGIR, CVPR, and ACM MM. His 15 techniques have been adopted into international video coding standards. His research interests include image and video compression, media communication, and media analysis and synthesis. He received the best paper award in the IEEE TCSVT 2009, the PCM 2008, and the SPIE VCIP 2007. He also received the IEEE Circuits and Systems Society 2012 Best Associate Editor Award. He serves as the TPC Chair of the MMSP 2011, the VCIP 2010, and the PCM 2009, and the special sessions Chair of the ICME 2010 and the ISCAS 2013. He also serves as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEM FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, and several other international journals.

Wen Gao (M'92–SM'05–F'09) received the Ph.D. degree in electronics engineering from The University of Tokyo, Japan, in 1991. He was a Professor of Computer Science with the Harbin Institute of Technology from 1991 to 1995, and a Professor with the Institute of Computing Technology, Chinese Academy of Sciences, from 1996 to 2006. He is currently a Professor of Computer Science with Peking University, Beijing, China.

He has authored extensively, including five books and over 600 technical articles in refereed journals and conference proceedings in the areas of image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interface, and bioinformatics.

Dr. Gao chaired a number of prestigious international conferences on multimedia and video signal processing, such as the IEEE ISCAS, the ICME, and the ACM Multimedia, and also served on the advisory and technical committees of numerous professional organizations. He served or serves on the Editorial Board for several journals, such as the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, the *Journal of Image Communications* (EURASIP), and the *Journal of Visual Communication and Image Representation*.